Thesis for doctoral degree (Ph.D.)
2018

Structural studies of proteins essential for fertilization

Elisa Dioguardi

Karolinska Institutet

From the Department of Bioscience and Nutrition
Karolinska Institutet, Stockholm, Sweden

STRUCTURAL STUDIES OF PROTEINS ESSENTIAL FOR FERTILIZATION

Elisa Dioguardi

Karolinska Institutet

Stockholm 2018

Published by Karolinska Institutet.
Printed by E-PRINT

ISBN 978-91-7831-132-3

Structural studies of proteins essential for fertilization

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Elisa Dioguardi

Principal Supervisor:
Prof. Luca Jovine
Karolinska Institutet
Department Biosciences and Nutrition

Co-supervisor(s):
Prof. Hans Hebert
Karolinska Institutet
Department Biosciences and Nutrition

Opponent:
Assoc. Prof. Stefania Gonfloni
University of Rome Tor Vergata
Department of Biology

Examination Board:
Assoc. Prof. Sebastian Deindl
University of Uppsala
Institute for Cell and Molecular Biology

Prof. Manuel Patarroyo
Karolinska Institutet
Department of Microbiology
Division of Tumor and Cell Biology

Prof. Rudolf Ladenstein
Karolinska Institutet
Department of Biosciences and Nutrition

to Elvira, who taught me how to be strong

to Piera, who encouraged me to become a scientist

ABSTRACT

Where do babies come from? Even though it sounds like a fairly obvious question, it took more than 200 years to discover that sperm, whose first observation dates back to the 17th century, can penetrate into the oocyte and fuse with it. Interestingly, this initial study was carried out in starfish, and since then, scientists have addressed the fascinating question of how gamete recognition functions using very different model systems.

Egg-sperm binding is a species-specific event, and it is thought to require binding proteins that are exposed on the surface on the egg and sperm. In this thesis, we focus our attention on how egg coat subunits are involved in this process and investigate it using X-ray crystallography and biochemistry.

The mammalian egg coat, also known as Zona Pellucida (ZP), is a thick extracellular matrix that surrounds the egg and is made of glycoproteins (known as ZP1-4) that assemble into long cross-linked filaments. Although ZP3 has been identified as a sperm receptor, it remains unclear as to how it interacts with sperm at molecular level. **Paper I** explores the potential role of sperm protein 56 (sp56) as a receptor for the egg coat protein ZP3. Its putative role as a ZP3-binding partner has been previously assessed by several *in vitro* assays using soluble proteins. Here, we express the recombinant form of hamster sp56 and attempt to reconstitute a hetero-complex with recombinant ZP3. Sp56 is secreted as a homo-hexamer held together by intramolecular disulfides. Considering the high avidity conferred by the homo-oligomeric state of the glycoprotein, the inability to observe any interaction to ZP3 by either pull-down or Far western blotting (Far WB) indicates that very-low affinity binding events might mediate egg-sperm interaction.

Different from the mammalian scenario, egg and sperm proteins in marine invertebrates are known to bind more tightly. **Paper II** sheds light on the event occurring when gametes first meet using abalone as a model system. We discover that egg coat functional units in mammals and invertebrates are structurally conserved and, remarkably, despite millions of years of evolution they rely on a common ZP-N fold domain to accomplish their functions. Crystallographic studies on individual subunits of the egg coat receptor VERL (vitelline envelope receptor for lysin) in complex with the sperm protein lysin, provides the first snapshot of egg-coat recognition and clarifies how species-specificity is maintained. By combining structural and biochemical data, the paper proposes the mechanistic model of how sperm penetrates the egg coat at the molecular level.

For structural studies of egg coat glycoproteins, it is essential to produce correctly folded material that is stable, highly homogeneous and is secreted in large amounts. To achieve this, we fuse glycoproteins to maltose binding protein (MBP) and describe in **Paper III** the advantages of such a method using mammalian and insect cell expression systems. The study herein shows how fusion to MBP boosts the expression of six challenging glycoproteins that are poorly secreted if unfused. MBP not only improves crystallization but provide initial phasing information for structural determination by molecular replacement.

Given the relevant role of egg coat subunits in fertilization, we investigate in **Paper IV** their functional relationship to two specular aspects of reproductive health. The study provides detailed information on how an anti-ZP-reagent recognizes its epitopes, and investigates how mutations in ZP genes relate to female infertility. By doing so, our findings further strengthen the importance of ZP in the development of novel contraceptive strategies and provide a better understanding of the impact of ZP mutations on human fertility

List of scientific papers

I. Dioguardi E, Han L, Jovine L (2018). **Molecular basis of egg-sperm interaction: characterization of hamster sp56.**
Manuscript

II. Raj I*, Sadat Al Hosseini H*, Dioguardi E, Nishimura K, Han L1, Villa A, de Sanctis D, Jovine L (2017). **Structural Basis of Egg Coat-Sperm Recognition at Fertilization.**
Cell; 169:1315-1326

III. Bokhove M*, Sadat Al Hosseini H*, Saito T*, Dioguardi E, Gegenschatz-Schmid K, Nishimura K, Raj I, de Sanctis D, Han L, Jovine L (2016). **Easy mammalian expression and crystallography of maltose-binding protein-fused to human proteins.**
J. Struct. Biol; 194(1):1-7.

IV. Dioguardi E*, Nishimura K*, Nishio S, Dietzel E, Inzunza J, Han L, Matsuda T and Jovine L (2018). **Molecular insights of ZP-based contraception and ZP-linked infertility.**
Manuscript

* These authors contributed equally to the work

CONTENTS

LIST OF ABBREVIATIONS

ZP	zona pellucida
ZP1-4	zona pellucida glycoprotein 1-4
sp56	sperm protein 56
CCP	sushi domains
VERL	vitelline envelope receptor lysin
MBP	maltose binding protein
PL	perivitelline space
VE	vitelline envelope
IHP/EHP	internal/external hydrophobic patch
CFCS	site consensus furin cleavage
TM	transmembrane domain
UMOD	tamm-Horsfall protein
EGF	epidermal growth factor
NTR	N-terminal regions
Far WB	Far western blotting
IVF	*in vitro* fertilization
ICSI	intracytoplasmic sperm injection
HTF	human tubal fluid medium
ART	assisted reproductive technology
PMSG	pregnant mare's serum gonadotropin
hCG	human chorionic gonadotropin
hGH	human growth hormone
C4bp	complement 4 binding proteins
HEK293	human kidney cells 293
CHO-K1	chinese hamster ovary

Sf9	*spodoptera frugiperda*
PNGase-F	peptide-N4-(N-acetyl-beta-glucosaminyl)
PPI	protein-protein interaction
Kd	dissociation constant
MST	microscale thermophoresis analysis
GlcNac	N-acetylglucosamine
Man	mannose
MS	mass spectrometry
TOF	time of flight
PDB	protein data bank
MIR	multiple isomorphous replacement
SIR	single isomorphous replacement
SAD	single wavelength anomalous dispersion
MAD	multi-wavelength anomalous diffraction
EM	electron microscopy
TEV	tobacco etch virus protease
IAM/OAM	inner acrosomal matrix/outer acrosomal matrix
SED-1	secreted protein containing EGF and discoidin/F5/8 complement domains
FSH	follicle stimulating hormone
TZP	transzonal projection

INTRODUCTION

"ex ovo omnia (all animals come from eggs)" - William Harvey, 1651

Fertilization is a crucial event for all sexually reproducing organisms. Knowledge of the molecular mechanism behind gametes recognition would represent a significant step forward in understanding how existing species are maintained and how new ones originate. More importantly, it has a high impact on human reproductive research in term of developing modern contraceptives or studying infertility.

This introduction compiles the most relevant literature of two well-characterized systems used for studying the mechanisms of fertilization; mouse (Study I and IV) and abalone, a marine mollusk (Study II). The final two subchapters give an overview on how the mammalian egg coat relates to contraception and infertility describing current applications (Study IV). Technical aspects concerning Study III are highlighted in the methodology section.

1.1 Fertilization: an overview

The two essential steps in reproduction are: meiosis that generates haploid gamete cells and, fertilization in which oocyte and sperm fuse to recreate a diploid cell, also known as zygote. Whether fertilization occurs externally (as in marine invertebrates) or internally (as in mammals), mammalian and non-mammalian oocytes present structural similarities. A thick extracellular matrix surrounds all eggs (Zona Pellucida, ZP, in mammals; perivitelline layer, PL, in birds; chorion in fish; vitelline envelope, VE, in amphibians and mollusks) and plays important roles in sustaining the oocytes during maturation, regulating species-specific binding to sperm and protecting the embryo prior to implantation[1]. Depending on the species, additional layers surrounding the extracellular matrix can be found as well, like cumulus cells in mammals or jelly layers in amphibians[2].

The origins of the extracellular matrix vary across animal classes, so that the egg coat components are either exclusively secreted by the oocyte, like in mice, hamsters and humans[3–5], or having multiple origins (oocytes and follicular cells) like in rabbits, dogs and pigs[6,7]. In birds, proteins of the PL are expressed either in the ovary by the oocytes and granulosa cells or in the liver and transported via bloodstream to the ovary[8,9].

Egg coat structures are more conserved among vertebrates and are made of up to seven ZP glycoproteins. For instance, mammalian genome contains three to four *ZP* genes (*Zp1*, *Zp2*, *Zp3*, and *Zp4*), whereas fish, *X. Laevis*., avian egg coats consist of four, five or six ZP proteins, respectively (Figure 1c).

Most of the progress in understanding the fertilization process comes from research on mice as well by using alternative models, like avian and non-invertebrate systems. By doing so, results have been extrapolated and applied to human fertilization which is often limited by the paucity of material and ethical restraints.

In mammalian fertilization, typically, very few ovulated eggs are in the oviducts, the site where fertilization takes place, with nearly 100 sperm that has to go through the cumulus cells to be able to reach the oocyte. Within the female reproductive tract, sperm reach its last phase of maturation called capacitation[10,11], a process that consists of a series of biochemical changes that enable cells to fertilize the egg. This includes flagellar hypermotility[12], expression of chemotactic responses and membrane remodeling. Once spermatozoa reach the oocyte and cumulus cells, they initiate the acrosomal reaction that leads to the release of ZP binding proteins. The acrosome is a unique secretory vesicle that is on the sperm head, and it releases enzymes that facilitate sperm to penetrate the egg coat[13]. A revised model of acrosome reaction, known also as “acrosomal exocytosis model”[14], discussed the existence of several intermediate states between acrosome intact and reacted sperm. Indeed, already during capacitation, acrosomal membranes start fusing and rapidly expose ZP-binding proteins (for example, sp56)[15]. Upon interaction with the egg coat, sperm progressively releases the acrosomal enzymes[16] whose role is, together with the hyperactivated motility, to facilitate the sperm penetration through the ZP[17].

Once spermatozoa reach the perivitelline space, it fuses with the oocyte. The egg protein Juno and the sperm protein Izumo[18] have been identified as key players for gamete recognition and, as shown by knockout mice studies, females that don’t express Juno are infertile, as well as males without Izumo[19].

To avoid polyspermy, lethal for the developing embryo, mammalian oocytes are thought to adopt several mechanisms including: 1) plasma membrane modifications leading to membrane depolarization[20], or removal of Juno receptors to avoid further fusion events (so-called “fast polyspermy block”); 2) or release of zinc, before cortical granules reaction initiates, as a mechanisms to alter the structure of ZP to decrease sperm penetration[21].

Ultimately, the cortical granules reaction is the only well documented event that enable oocytes to avoid polyspermy[22] (also known as “slow polyspermy block” or “ZP-hardening”). Immediately after fertilization, cortical granules - membrane bound organelle that locate along the cortex of the oocyte - release glycosidases that target carbohydrates of the egg coat together with metallo-proteases that cleave the ZP subunit ZP2 (ZP2f)[23]. Several lines of evidence showed that conversion of ZP2 to ZP2f relates to a more compact and hard ZP that is less permissive to sperm penetration[23,24].

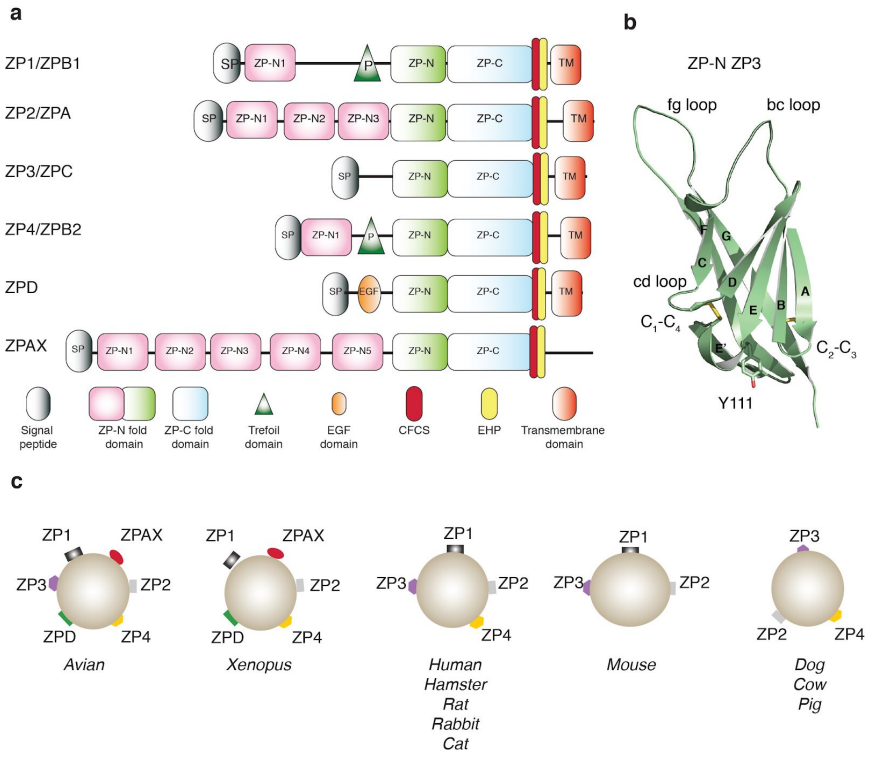

Figure 1. ZP proteins and egg coat composition in vertebrates. Modified picture from *Goudet et al* 2008 and *Monnè et al* 2008. a) Domain architecture of ZP proteins; ZP-N domain (pink); trefoil domain (dark green); ZP-N of ZP module (purple); ZP-C domain (cyan); furin cleavage site (CFCS; red) and external hydrophobic patch (EHP; yellow); EGF domain (orange) and TM (red) b) Cartoon representation of ZP-N domain. Disulfides are shown in stick and conserved tyrosine is marked. c) Egg coat composition in different types of vertebrates.

1.2 Mammalian zona pellucida

The ZP appears in primordial follicles between the oocyte and the surrounding granulosa cells[4,5,25,26] and it is separated from the oocyte by the perivitelline space. It is a fibrous protein matrix that progressively increases its thickness to several µm width during folliculogenesis. Depending on the species, thickness can vary from 7 µm (mouse) to 24 µm (pig). Although being physically separated from granulosa cells, gap junctions between

the ZP and somatic cells ensure transport of nutrients to the oocyte throughout the follicular development[27].

In mammals, ZP filaments are made of 3-4 glycoproteins that assemble through a common C-terminal ZP module (Figure 1,2) - a 260 amino acid module made of two domains (ZP-N and ZP-C) that contain 8/10 invariant cysteines[28,29].
As a general accepted model, ZP2, ZP3 and ZP4 incorporates into long filaments that are covalently cross-linked by ZP1 homodimers[30]. Whereas ZP2 and ZP3 are important for egg coat assembly and to regulate sperm-binding[31–34], the function of ZP4 remains still unknown.

The current view suggests that filament formation occurs outside of the cell. Glycoproteins are translocated to the membrane (via TM) and are proteolytically cleaved at the consensus furin site (CFCS; R-XX-R) which locates between the ZP module and external hydrophobic patch (EHP; β-strand G) (Figure 1a). Upon EHP release, the internal hydrophobic patch (IHP; β-strand A) becomes exposed and triggers filament growth[29].
The first crystallographic structure of the ZP-N domain of mouse ZP3 gave better insights of the structural organization of the egg coat and revealed that other egg coat subunits, mammalian and not, contain additional copies of ZP-N[35,36]. At first sight, the ZP-N domain resemble an Immunoglobulin (Ig) domain in which the two antiparallel β-sheets are made of four (A-B-E-D; β-sheet 1) or three (C/E′-F-G; β-sheet 2) β-strands, respectively. Both side are held together by two conserved disulfide bonds presenting the connectivity C1-C4 and C2-C3. As additional features, ZP-N is characterized by having a E′ β-strand in the β-sheet 2 and an invariant tyrosine in the β-strand F (Figuren 1b). Analysis of the full-length structure helped to clarify the dual function of ZP3, both as a building block for filaments formation and as a sperm receptor. Considering that ZP3 is prompt to form filaments, it was essential to perform crystallographic studies by working with a highly stable protein to avoid protein aggregation. This was achieved by expressing ZP3 in its uncleaved form, with the EHP retained inside the ZP-C and without the TM[37]. The overall structure shows ZP3 engaged in a homodimer conformation with ZP-N and ZP-C sharing a common ZP-N fold despite their very low sequence similarities. These two domains are spaced by an extended amino acid linker that contains an O-glycan involved in sperm binding (T165). Both domains hydrophobically interact through the EHP which constitutes the G β-strand of ZP-C.

A numbered of other polymerizing human proteins have been discovered to contain a ZP module[38]. For instance, Tamm-Horsfall protein uromodulin (UMOD) that makes homopolymers in the urine or α- and β-tectorin that constitute the tectorial membrane of the cochlea. Structural comparisons between ZP module of ZP3 and UMOD[37-39] helped to elucidate how polymerization might take place within the egg coat[30]. Unlike ZP3, ZP module of UMOD structure has a more elongated shape with ZP-N and ZP-C positioned apart by a long structured linker that contains an α-helix and β-strand (Figure 2a). As a result of this conformation, the two ZP-Ns interact and trigger polymerization of UMOD. The model suggests that the interdomain linker between ZP-N and ZP-C dictates the type of polymerization and depending on whether it is structured or not, it either leads to homopolymers (UMOD) or to heteropolymers (ZP3, ZP2 and ZP4).

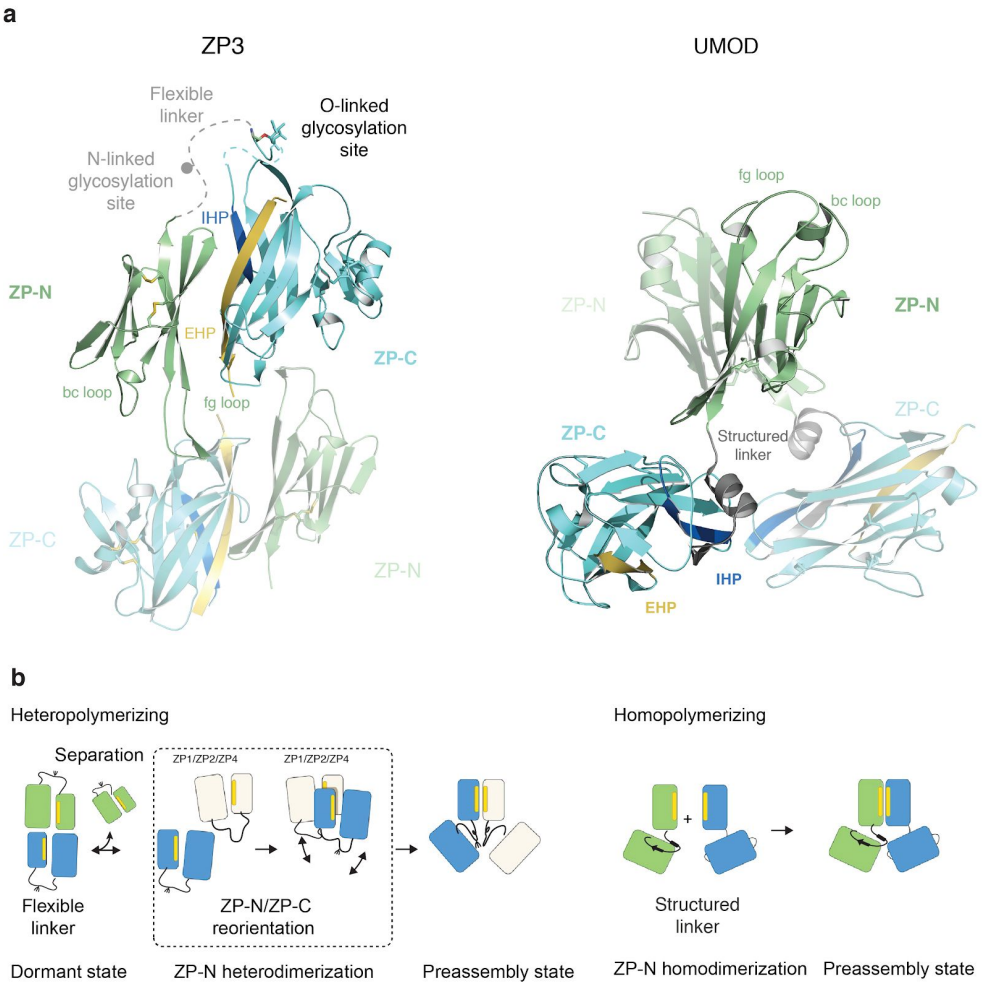

Figure 2. Polymerization mechanism of ZP module. Modified picture from *Bockhove and Jovine (2018)*. a) Structures of ZP module of ZP3 and UMOD showing the arrangement of the dimer. ZP-N (green), ZP-C (cyan); IHP and EHP are marked with different colors. b) Proposed mechanism of polymerization through ZP module.

1.3 Zona Pellucida proteins ZP1 and ZP4

ZP1 and ZP4 have arisen from a common gene as a result of a gene duplication event that occurred after the lineage of fish[40]. As a consequence, they share sequence similarities and present a similar domain organization (Figure 1a).

Depending on whether ZP1 or ZP4 is a pseudogene, eutherian ZPs can be divided into three subgroups:

1) species with ZP made of three ZP subunits with ZP4 pseudogene, to date only house mouse;

2) species with ZP made of three subunits with ZP1 pseudogene, like pig, cow, dog and fox;

3) species with ZP made of four subunits like human, rat, cat and hamster[41].

In mice, ZP2 and ZP3 are the most abundant glycoproteins and they account for roughly 90% of the total protein content of the egg coat. On the other hand, ZP1 (~120 kDa; human) is less expressed and its role is to covalently cross-link different ZP filament to give rise to a 3D matrix[42]. Accordingly, female knockout mice lacking *Zp1* produced oocytes with a thinner and poorly organized matrix[43]. Subfertility was due to precocious hatching of the eggs and early embryonic loss.

Although ZP4 has been also thought to function like ZP1, no current data sustain this hypothesis.

ZP4 (~65 kDa; human) was first discovered in human oocytes by mass spectrometry analysis[44] (MS) and further in a number of other species. Considering that the human spermatozoa do not bind to the mouse ZP, it was initially thought that ZP4 might had been involved in egg coat-sperm binding in humans. However, oocytes of transgenic mice expressing mouse ZP1-3 and human ZP4 failed at binding human sperm[45,46].

1.4 Zona Pellucida protein ZP3

ZP subunits are heavily post-translationally modified with N-linked and O-linked carbohydrates whose compositions, mainly sialylation and sulfation, have been shown to vary among several mammalian species[47]. ZP3 (~55 kDa in mouse) is a glycoprotein essential for filament formation, as exemplified by the phenotype of *Zp3* null mice ($ZP3^{-/-}$) whose females failed to produce ZP and were therefore completely infertile[32].

Most of the work to assess mouse ZP3 function was performed by *in vitro* assays using solubilized purified glycoproteins. It was discovered that:

1) among all the three subunits, mouse ZP3 purified from unfertilized eggs inhibited sperm binding to the egg coat at a nanomolar concentration (so-called *competition assay*), while the glycoprotein purified from 2-cell embryos failed[48];
2) glycopeptides of mouse ZP3 generated by extensive proteolysis retained the ability of inhibiting sperm binding, albeit at higher concentration (50 fold) than the one used for full-length protein[49];
3) gold-labeled ZP3 was found to bind to acrosome-intact sperm[50];
4) beads covalently coated with ZP3 could pull-down sperm with an intact acrosome[51];
5) ZP-sperm binding events involved carbohydrates. Nonetheless, protein–protein interactions must be considered as well, because it is likely that ZP proteins adopt particular conformations that optimize the presentation of glycans for recognition by sperm. It was initially hypothesized that O-linked attached to S332/334 might have been implicated: however, their removal in knockout studies clearly showed that they had no impact in egg-sperm binding and on animal fertility[24]. Remarkably, analysis of the crystal structure of avian ZP3 revealed extra electron density corresponding to a O-linked glycan attached to T168 which is part of a motif highly conserved in mammals. Mutation of this residue decreased chicken sperm binding to ZP3 to 20%[37]. Consistent with the important role of glycosylation in egg-sperm recognition, a recent study showed that synthetic sialyl-Lewis^x-Lewis^x (SLe^x–Le^x), which represents the most abundant type of carbohydrate within human oocytes, inhibited human sperm binding to the ZP[52];
6) ZP3 purified from unfertilized egg induced sperm to complete acrosome reaction *in vitro* to the same extent of the calcium ionophore, while the glycoprotein purified from embryos couldn't[53];
7) binding to ZP3 activated G-proteins[54] and voltage Ca2+ channels[55] that increase intracellular Ca2+ levels which, in turn initiate the acrosome reaction.
All these shreds of evidences merged in a first model which suggested ZP3 as primary receptor for sperm and inducer of the acrosome reaction, while ZP2 role was to keep sperm loosely bound to the ZP following acrosome reaction (secondary receptor).

Advances in genomic manipulation and generation of transgenic/knockout mice gave the opportunity to further study functions of ZP proteins within the egg coat matrix. By doing so, a new mechanism for fertilization called the "supra-molecular structure model" was proposed. This new model discarded the importance of ZP3 and acrosome reaction. Instead, it suggested that the overall ZP, whose structure changes upon post-fertilization cleavage of ZP2, regulates sperm binding.

1.5 Zona Pellucida protein ZP2

Crystallographic studies on isolated ZP3 ZP-N by our group revealed that ZP2, as other egg coat glycoproteins involved in fertilization[35], contains additional ZP-N copies in the N-terminal region (NTRs) (Figure 1a). Post-fertilization ZP2 cleavage identified within the NTR was shown to induce structural changes of the ZP, conferring mechanical stiffness and increased resistance to dissolution by different proteases (ZP hardening)[56]. Kinetic studies of the conversion of ZP2 to ZP2f in mice showed that the reaction occurs quite fast and already 5 minutes after egg activation 50% of ZP2 is proteolyzed[56]. The temporal window of the ZP hardening was discovered to vary across mammals leading to specific properties of the egg. Nevertheless, what is common to all the species is that when ZP is more compacted, sperm fails to bind and penetrate the egg coat. According to the current view of how ZP filaments are organized, the ZP module constitutes the filament core, whereas the NTRs are likely to project out. Upon ZP2 cleavage, NTRs belonging to different filaments are thought to interact, resulting in overall hardening of the ZP[36].

Ovastacin is a metalloprotease released by cortical granules that is responsible for ZP2 cleavage[57] which occurs at the site 166-LA↓DE-170 in mouse[23]. This generates two fragments of approximately 90 kDa and 30 kDa that remain linked by a single disulfide bond. Interestingly, western blot analysis using a monoclonal antibody targeting the first domain of ZP2 showed that ZP from 2 cell-embryos contains multiple fragments, rather than just the 30 kDa, suggesting that further proteolysis events occur as well. Ablation of the ovastacin gene ($Ast^{-/-}$) compromised ZP2 processing and gave rise to a similar phenotype of the one observed in mice containing the uncleavable form of ZP2 (huZP2 rescue and $ZP2^{Mut}$) where sperm binding endured until the 2-cell stage[24,58]. In both cases animals were subfertile due to pre-embryonic loss. Interestingly, unlike the $Ast^{-/-}$ null mice, subfertility of $ZP2^{Mut}$ transgenic mice was shown to be more pronounced due to a thinner ZP (4 µm instead of 7 µm in normal mice). Small releases of ovastacin were discovered to occur spontaneously both *in vitro* and *in vivo* causing pre-hardening of the ZP before fertilization. A recently proposed mechanism has described the serum protein fetuin-B as inhibitor for ovastacin to prevent early cleavage events and precocious ZP hardening[59].

The ZP2 function has been explored using transgenic mice. This type of approach gave the opportunity to elucidate the role of ZP subunits in an engineered ZP by performing gain-of-function assays or mapping regions involved in gamete recognition. In this regard, human sperm was found to bind ZP of mouse oocytes expressing human ZP2[34,60]. In

addition, beads coated with recombinant constructs encompassing the N-terminal region of ZP2 could pull-down sperm[34,61].

This last set of experiments defined a different model for fertilization based on ZP2, known as supra-molecular structure model. A recent analysis hypothesized that both mechanisms, the first one described by Wassarman and colleagues and this second one might indeed contain common elements and could be merged together (Figure 3). According to this "hybrid model", post-fertilization cortical reaction would inactivate the ability of ZP2 to bind acrosome-reacted sperm and simultaneously affect the presentation of ZP3 carbohydrate epitopes that mediate the first egg-sperm recognition[62].

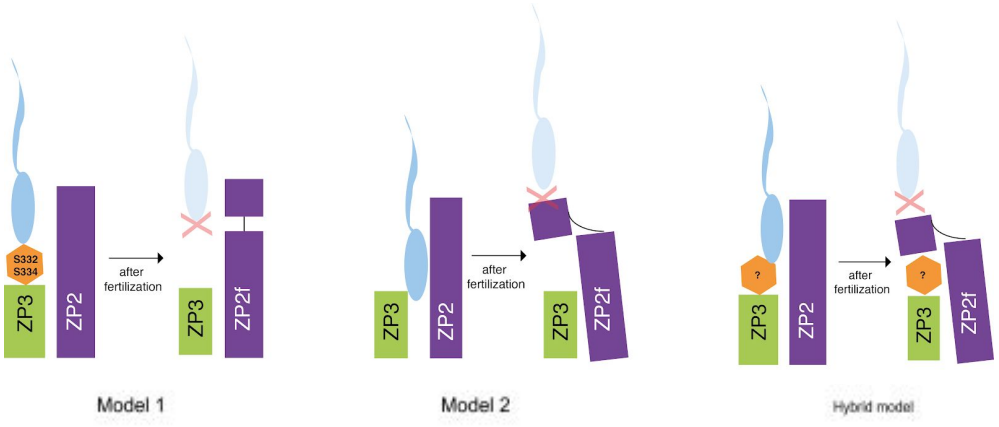

Figure 3. Model for sperm binding taken from *Visconti and Florman* (2014).

1.6 Transgenic mouse models

Knockout studies contributed with valuable insights to understand the function of ZP proteins and their impact on fertility. The most severe phenotype was the one observed in ZP3-deficient mice ($Zp3^{-/-}$) that produced zona-free oocytes and were completely infertile[32]. The absence of the egg coat impaired the correct development of the oocyte as evidenced by reduced number of oocytes that reached pre-antral follicle stage. Besides, all post-ovulatory follicles appeared visibly disorganized due to the inability of granulosa cells to tighten around the egg. IVF analysis of $ZP3^{-/-}$ free zona-oocyte showed that these eggs completed the first meiosis *in vitro*[63] and, could be fertilized *in vitro* (Table 1). Although they progressed to blastocyst stage, no live births were identified after transfer due to precocious embryo loss[31]. Mice with heterozygous insertional mutation in Zp3 gene ($ZP3^{+/-}$) produced oocytes with a thinner ZP and were fertile as controls. Similar to $Zp3^{-/-}$ phenotype, ablation of ZP2 ($Zp2^{-/-}$) caused infertility[31] due to absence of the ZP around ovulated oocytes (Table 1). Although a very thin and fragile ZP (made of ZP1 and ZP3) was initially detected in large antral $Zp2^{-/-}$ follicles, it disappeared after ovulation. Instead,

oocytes that do not express ZP1 form a loose matrix made of ZP2 and ZP3 and mice are subfertile.

	Normal	*Zp1* null	*Zp2* Null	*Zp3* null	Zona-free normal
Fertilized eggs	518 (100%)	452 (100%)	227 (100%)	395 (100%)	395 (100%)
Two-cell embryos*	417 (81%)	337 (75%)	120 (53%)	196 (50%)	179 (45%)
Blastocysts*	253 (61%)	170 (50%)	36 (30%)	81 (41%)	72 (40%)
Transferred‡	164 (100%)	162 (100%)	31 (100%)	43 (100%)	56 (100%)
Live births§	24 (15%)	44 (27%)	0 (0%)	0 (0%)	5 (9%)

*Number of embryos (percent that progress from previous stage).
‡Number of cultured blastocysts transferred to pseudopregnant females.
§Number of live births (percent of transferred blastocysts).

Table 1. Table taken from *Rankin et al (2001)* showing IVF and embryogenesis of $Zp1^{-/-}$, $Zp2^{-/-}$ and $Zp3^{-/-}$ oocytes.

Transgenic mouse studies were performed using both human and mouse sperm in order to understand which one of the glycoproteins is crucial for egg-sperm interaction. The table that follows (Table 2) aims at summarizing our current knowledge derived from the transgenic mouse lines developed for studying the impact of ZP glycoproteins to fertilization. Aspects like egg-sperm binding, *in vitro* fertilization and animal fertility are highlighted. In some cases, additional notes are presented as well.

Although some of the questions remain unsolved, for instance, the status of the acrosome or the exact sequence of events, the overall picture shows that recognition mostly depends on the supra-molecular structure of the ZP which may be a consequence of the cleavage status of ZP2.

ZP contains	Sperm-binding to ZP	IVF	Fecundity	Message of the paper	Notes
mZP1, huZP2, mZP3 [58,60]	Mouse sperm bind to ZP of unfertilized eggs and embryos	IVF performed with mouse sperm (reduced success rate). 1 hours after insemination mouse sperm had intact acrosome	Mice are subfertile and ovulate less / Thin ZP causes pre-embryonic loss	Mouse ovastacin does not cleave huZP2 and phenotype looks similar to the one of $ZP2^{Mut.}$ ZP doesn't induce acrosome reaction	*Notably, after 1 - hour the ZP post fertili zation block has been already initiated. It is expectable to see sperm acrosome intact. So, it would be interesting evaluating acrosome reaction in the early minutes.*
mZP1, $mZP2^{Mut}$, mZP3 [24] mZP2 cannot be cleaved	Mouse sperm bind to ZP of unfertilized eggs and embryos	IVF/ Oocyte can be fertilized and developed to embryo; No sperm in the perivitelline space	Mice are subfertile and ovulate less / Thin ZP causes pre-embryonic loss	ZP2 cleavage makes the ZP permissive to sperm binding	
mZP1, mZP2, huZP3 [34,58]	Mouse sperm bind/ Human sperm don't bind	IVF/ Oocyte can be fertilized	Mice are subfertile and have less ovulated eggs / Thin ZP causes pre-embryonic loss	huZP3 does not rescue binding of human sperm	*ZP3 glycosylation is involved in sperm adhesion. Maybe huZP3 expressed by mouse oocytes does not contain the same type of glycans of wild type.*
mZP1, mZP2, $mZP3^{Mut}$ [24] mZP3 S332A and S334	Mouse sperm bind to ZP	IVF/Oocytes can be fertilized	Mice are fertile like controls	O-linked glycan attached to S332 and S334 do not play a role in ZP-sperm binding	*Other O-linked glycan might be involved, for instance T155 in mouse*
mZP1, huZP2, mZP3 [34]	Human sperm bind to ZP when it is capacitated longer than 4 hours	-	Mice are subfertile	Presence of huZP2 is required for human sperm binding	*Human sperm-binding after fertilization was not examined. Does un-cleaved huZP2 give a phenotype similar to mZP^{Mut}?*
mZP1, mZP3, huZP4 (moQUAD-$ZP2^{Null}$) [46]	Mouse sperm don't bind	-	Mice are sterile and have less ovulated eggs / no embryos detected	moZP2 is essential for sperm binding *in vitro* and *in vivo*	*huZP4 increased thickness of the ZP/ Since IVF was not performed, we don't have information about the quality of the oocyte.*
mZP1, mZP3, $mZP2^{Trunc}$, huZP4 [46]	Mouse sperm don't bind	-	Mice are sterile and have more ovulated oocytes than moQUAD-$ZP2^{Null}$ / no embryos detected	The N-terminus (residue 51-149) of ZP2 is necessary for sperm binding	
huZP1, huZP2, huZP3, huZP4 (huQUAD) [34]	Mouse sperm bind. Human sperm bind to ZP of unfertilized eggs, but not the one of embryos obtained after IVF with mouse sperm	IVF/Oocyte can be fertilized *in vitro* using mouse sperm; human sperm penetrate the transgenic ZP and accumulate in the perivitelline space (all acrosome-reacted)	Mice are 50% subfertile due to abnormal oocytes development	Human sperm bind, penetrate the mouse eggs with humanized ZP/ Sperm bind to ZP2	*Previous work showed that mouse ovastacin doesn't cleave huZP2. Therefore, human sperm should keep binding the ZP of embryos (data not shown).*
huZP1, huZP3, huZP4 [46]	Human sperm don't bind	-	Mice are sterile. Although mice could ovulate, no embryos are detected	Human sperm binding depends on huZP2	*Since IVF was not performed, we are unsure whether infertility is caused by developmental abnormalities or by the absence of binding.*
huZP1, mZP2, huZP3, huZP4 [46]	Human sperm don't bind	-	mZP2 rescue mouse fertility	mZP2 and huZP2 are necessary for sperm binding	*Does mouse sperm bind to these eggs? Does human sperm bind to the ZP of embryos?*

Table 2. Summary of transgenic mice developed for studying functions of ZP proteins

1.7 How does sperm bind to the egg coat? Sp56 as putative receptor for ZP3

Interspecies fertilization is generally avoided and the egg coat is the major barrier that prevents it. Despite some exceptions, for instance mouse sperm is able to bind to human ZP *in vitro* or human sperm recognizes oocytes of other primates[64], gamete recognition involves binding of highly species-specific proteins.

Sperm is much smaller (roughly 10000 times) than the egg and is suggested to expose egg-binding proteins on its head that help sperm to bind and penetrate the egg coat. There are three important anatomical features in this cell to keep in mind: the head where the nucleus is surrounded by a specialized vesicle called the “acrosome”; the midpiece full of mitochondria and the long tail essential for propulsion.

The development of IVF techniques and progress in understanding the mechanisms behind sperm capacitation have contributed to clarify some molecular aspects of fertilization in mammals. IVF became a commonly used assay to test functions of individual proteins in gamete recognition. By using very different types of biochemical approaches, several candidates have been described in different mammalian species, mostly in mouse though, that possess egg coat binding activity (zonadhesin, SED-1, spermadhesin, β-galactosyltransferase)[65]. However, their role remained dubious as homozygous null male mice did not show evident defects at the level of egg-sperm binding and were fertile.

One of the sperm proteins that was discovered to have a ZP3 binding activity was sp56[66]. Among all the candidates, sp56 showed all the requirements to be the truly receptor. It is an intra acrosomal protein[67] identified by photoaffinity cross-linking and discovered to bind mouse ZP3 *in vitro*[68,69].

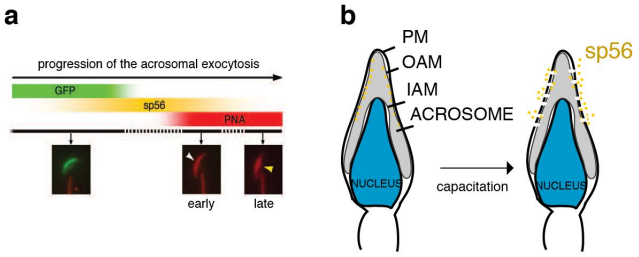

Figure 4. Release of sp56 during acrosome reaction. Modified picture from *Hirohashi et al* (2011) and *Wassarman* (2009)

Sp56 was only found in acrosome-intact sperm and was discovered to bind a galactose-affinity column, a type of glycan that is functionally important in ZP3 during sperm adhesion[70]. After capacitation, pH and intracellular Ca2+ levels transiently increase and trigger multiple fusion events between the outer (OAM) and internal membranes (IAM) of the acrosome. As a result, sp56 becomes exposed to the external environment and available for ZP3 binding. Accordingly, *in vivo* imaging of EGFP-mouse sperm during IVF showed a progressive exposure of the glycoprotein during acrosome exocytosis (Figure 5)[71] . To date, only mouse, rat and guinea pig sp56 have been characterized. Based on its primary sequence, the glycoprotein was discovered to belong to the same family of C4 binding proteins (C4bp)[15]. This protein, as well sp56, is characterized by a repetition of small 60 residues domains, called Sushi domain (CCP), each of them containing four conserved cysteines. The C-terminal region of C4bp contains a common oligomerization domain that is responsible for assembly of the protein into a 500 kDa homo-oligomeric complex (in the case of mouse sp56 six or more monomers), further stabilized by intermolecular disulfides[72]. This conformation was shown to confer to the protein high stability and flexibility[72].

Recombinant mouse sp56 was expressed as properly folded glycoprotein using human kidney cells 293 (HEK293) and used for several functional assays. More specifically, it was shown that soluble recombinant sp56 inhibited mouse fertilization *in vitro*[73] and beads coupled with the protein covered the unfertilized egg and not the 2-cell embryos[74].

1.8 Abalone as a model system: VERL and lysin

Unlike mammals, the egg coat and sperm proteins of abalone are known to interact and bind each other tightly. Abalone is a marine gastropod mollusk of the family *Haliotidae.* They have separate sexes that spawn in the same areas releasing million of gametes. Thank to the abundance of their gametes that are easy to harvest, abalone become an optimal model system to investigate the events behind egg-sperm recognition.

When sperm is released in seawater, it is chemotactically attracted by the egg and after having penetrated the jelly coat, it releases the acrosomal content. Lysin is the most abundant protein of the acrosome together with a protein of 18 kD. It creates a hole in the VE by splaying apart egg coat filaments in a non-enzymatic manner[75] (Figure 5a). Notably, a non-enzymatic mechanism facilitate sperm penetration without compromising the entire structure of the VE that has to protect the developing embryo.

Currently, there are circa 60 abalone species leaving in different habitats and hybrids are rare[76] in nature. Accordingly, *in vitro* VE dissolution assays showed that it takes 10 times more amount of sperm to fertilize eggs when species are different[77,78] (Figure 5c). This would suggest the existence of species-specific binding proteins on the egg that limit cross-fertilization and allow sperm from the same species to fertilize first.

Lysin is a 16 kDa protein that contains a high number of positively charged residues. It is not glycosylated and have no cysteines. Comparisons of multiple lysin sequences from different species shows a high degree of similarities except for the N and C-termini that are more divergent and later on suggested to be involved in modulating the binding to the VE filaments. Among the proteins involved in fertilization, lysin was discovered to evolve up to 50 times faster than proteins in mouse and humans[79]. By definition, rapidly evolving proteins experienced a high pace of amino acid substitutions between species[80] and this can be calculated by looking at the ratio between non-synonymous/synonymous substitutions; values greater than 1 would indicate positive Darwinian selection.
The overall crystal structure showed lysin adopting a 5 helix-bundle domain[81] with the N-terminal (1-12) projecting out of the domain. The two prominent structural features are characterized by a long 16-residues hydrophobic patch containing hydrophobic and aromatic residues that stick outside of the protein giving unusual hydrophobic properties; and secondly, the presence of a large positively charged area that contributed to increase the isoelectric point of the protein to 10.

The abalone egg coat is made of more than 30 different proteins[82] that have in common the C-terminal ZP module. To discover the binding partner of lysin, dissolved VE was passed over a column coupled to lysin[77] and elution was then analyzed. VERL (vitelline receptor for lysin) is a massive protein (~1 MDa), representing the most abundant component of the VE and it is heavily glycosylated (glycan content accounts for 50% of the total mass). To confirm the function as receptor for lysin, isolated VERL molecules were used in an inhibitory competition assay and were shown to affect the ability of lysin to dissolved isolated VEs[77]. Already by looking at the primary sequence of VERL, a large number of negatively charged residues, possibly complementary to lysin, were identified. The glycoprotein presents a C-terminal ZP module and a repetition of 22 tandem repeats in the N-terminal, hypothesized to fold like ZP-N[83] (Figure 5d). Interestingly, the first two repeat were discovered to be under positive selection and their sequence diverged more than the remaining identical 20 repeats. Biophysical characterizations (dynamic light

scattering, chemical cross-linking and FRET analysis)[84,85] showed that lysin exist in solution as a dimer and when in contact with VERL it becomes a monomer.

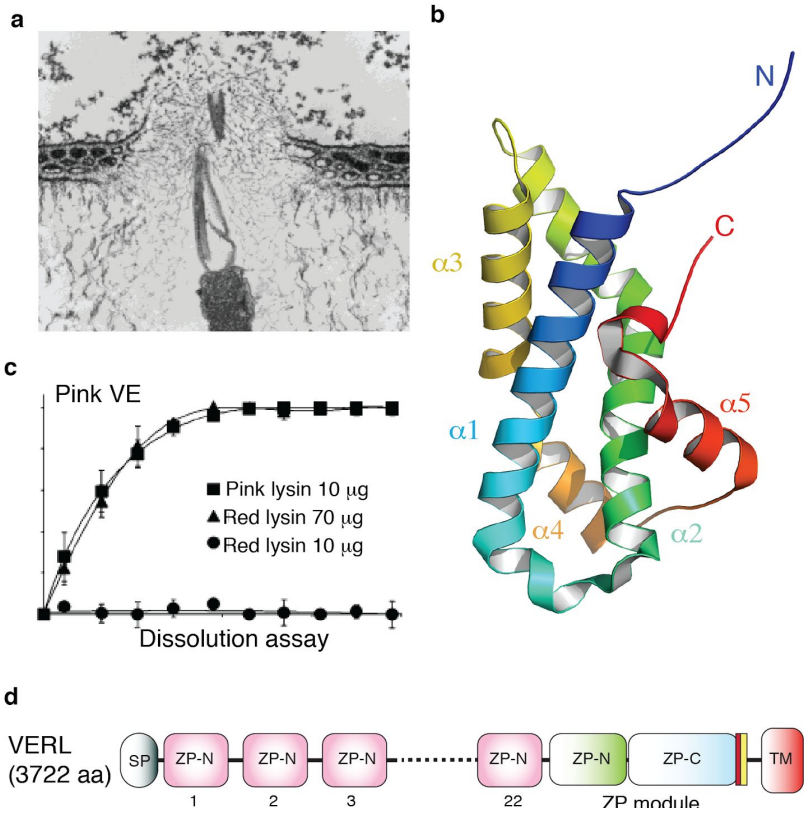

Figure 5. VERL and lysin. a) Electron microscopy picture showing abalone sperm making a hole in the VE. Picture taken from *Lewis C. et al* (1982). b) Cartoon representation of red lysin colored rainbow (modified picture from *Shaw et al* (1993); 1LIS). c) VE dissolution assays of two different abalone species showing that it is necessary adding much more red lysin to dissolve VE of pink abalone (modified picture from *Lyon and Vacquier* (1999). d) Domain architecture of VERL (modified picture from *Swanson W. et al* (2011)). Signal peptide (SP); ZP-N (pink); ZP-N of ZP module (green); ZP-C (cyan); EHP and CFCS (red and yellow respectively); transmembrane domain (red).

1.9 Implications of ZP for human infertility

In humans, oocyte development begins during fetal life when oogonia proliferate in the ovary and develops into primary oocytes (Figure 6a). At birth they are around 300000 and they are arrested in prophase I of meiosis I until puberty comes[86]. At this stage primary oocytes are surrounded by a flat layer of granulosa cells. After menarca, periodically a group of primordial oocytes leave the resting pool to enter in a grow phase where their size increase and granulosa cells proliferate from single to a multilayered ring. The transition is a slow process, independent from follicular stimulating hormone (FSH) whose receptors are expressed in granulosa cells during the transition from secondary to pre-antral follicles.

During this growth phase oocytes complete their maturation reaching metaphase II where they arrest again until fertilization occurs (See Figure 6a). The ZP appears in primary follicles and it is indispensable to ensure correct cellular connections during folliculogenesis (transzonal projection; TZPs) between the nursing granulosa cells and the oocyte[87]. Structural abnormalities of the ZP impair cellular communication with the cumulus cells and thus affect the quality of the oocyte.

As the oocyte increases its size, the ZP grows from inside to outside[88] reaching several micrometer thickness. According to studies carried in mice and hamsters, ZP genes are coordinately expressed throughout the whole folliculogenesis process[4,89].

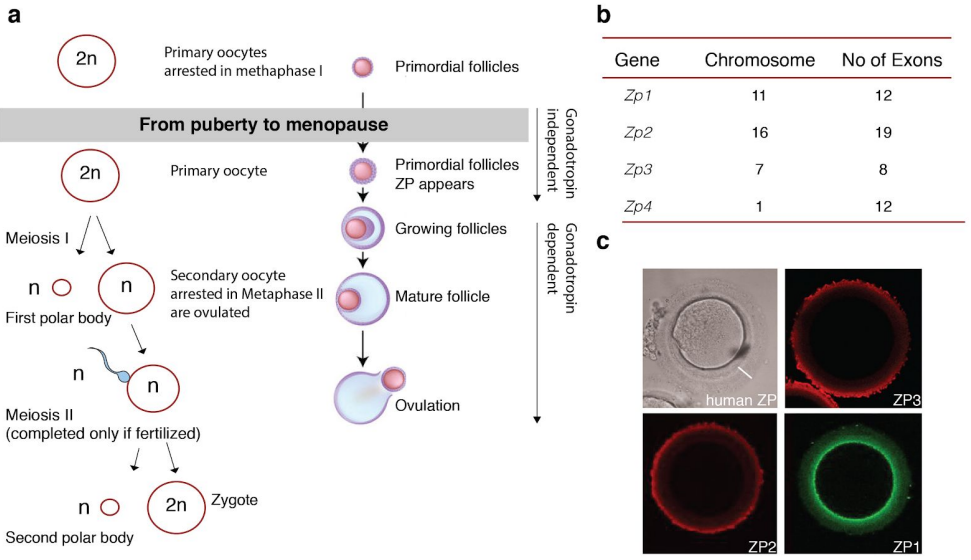

Figure 6. Folliculogenesis and oocyte development. a) Meiosis I and II taken from - Female Reproductive system; Physiology -. b) Genes of human ZP proteins. c) Immunostaining of ZP of a human oocyte taken from *Dai et al* (2018). Staining against ZP4 is missing.

Assisted reproductive technology (ART) has provided valuable insights of some of the properties of the oocyte whose quality is currently the limiting factor for female infertility. Thus, a way to select a good oocyte is to characterize the morphologically aspects of the ZP which is considered a good indicator for oocyte development. It is known that thickness may vary among individuals, but significantly increases with age[90] and in smokers[91]. Some studies showed that thick ZPs gave poor fertilization IVF rates and had a negative impact during blastocyst expansion, whereas a thinner ZP favors embryo hatching[92] and gives better fertilization rates[93].

Scanning electron microscopy revealed the ZP as a multilaminar structure. Analysis carried with polarized light microscopy is currently used as non-invasive technique to select optimal oocytes by visualizing the structural organization of the inner layer. More

specifically, a highly birefringent signal has been associated to high fertilization rate and pregnancy outcome[94] while a low birefringent ZP correlates to miscarriage[95]. However, it is still a matter of debate whether the ZP and ART success are significantly related as shown by several other contradictory studies[96]. An additional biomarker for embryo quality which is gaining more attention, is the zink sparks, a massive release of zinc ions from the oocyte after fertilization[97]. Retrospective studies showed that flux of zinc is more pronounced in those eggs that develop into blastocyst[98].

Defective egg coat-sperm binding is one of the reasons for failure of conventional IVF treatments[99] which is estimated to be around 10-15%. Both sperm or oocyte defects are involved in this process and even though it can be bypassed by intracytoplasmatic sperm injection (ICSI), understanding how female factors contribute to egg-sperm binding would constitute an important tool to select ART strategies. Initial studies investigating the relationship between genetic factors and unexplained female infertility[100,101,102] failed at finding a definitive link to infertility due to the low number of affected patients and the inability to draw good statistics. However, a 1.5 fold increase of mutations in Zp1 and Zp3 genes[102] have been observed, suggesting an important role of these two glycoproteins in human fertilization.

Currently, an increased number of novel ZP gene variants[103–107] discovered in patients presenting visible defects of the ZP has been under examination (Table 3).

Exploring the genetic setting and understanding whether ZP mutations have an impact on female infertility might help to evaluate optimal treatments and maximize the outcomes to achieve pregnancy.

	Variant	Inheritance	Oocyte	Publication
Zp3	A134T	Heterozygous	Oocytes with no ZP	[106]
	G31R	-	No link to oocyte lysis	[102,101]
	S264P	-	No link to ZP abnormalities	[100]
	K247K	-	No link to ZP abnormalities	[100]
	A183T	-	-	[102]
Zp2	C557H-fr5*	Homozygous	Thin ZP that lacks ZP2	[107]
	C557W-fr5*	Homozygous	Thin ZP that lacks ZP2	[107]
	L698* (no TM)	Heterozygous	Thin ZP	[105]
	C557R	Heterozygous	Degenerate oocytes and	[103]
	R524S	Heterozygous	Crackled ZP	[103]
	G36V	-	No link to ZP abnormalities	[100]
	D173D	-		[100]
	P249P	-		[100]
Zp1	W83R	Heterozygous	Degenerate oocytes and	[103]
	W471*	Heterozygous	Crackled ZP	[103]
	P360T	-	No link to oocyte lysis	[101]
	S389-fr15*	Homozygous	Zona-free oocytes	[104]
	V19V	-	No link to ZP abnormalities	[100]
	T158I	-	No link to ZP abnormalities	[100]
	V289V	-	No link to ZP abnormalities	[100]
	V93I	-	-	[102]
	I158T	-	-	[102]

Table 3. Non-synonymous and missense mutations reported in *ZP* genes. Sequence variations of non-coding and coding regions (synonymous mutation) found in *Zp1, Zp2 , Zp3* and *Zp4* are not reported in this table. Empty boxes (marked with -) indicate that the type of inheritance or oocyte was not described.

1.10 Application of the ZP: female contraception

To date there are multiple methods available to avoid pregnancy. Together with condom and female sterilization, hormonal contraceptives are the most popular anticonceptional methods, for instance in USA[108]. All formulations contain estrogen or progestins or both that inhibit secretion of gonadotropins by introducing a negative feedback and thus suppress ovulation. The list of potential side effects is long in which major ones are nausea, weight gain, chloasma, decreased libido, breast tenderness. According to the world health organization (WHO), they are also risk factors for developing cardiovascular complications, breast and cervical cancer, hepatic adenoma[109].

Immunocontraception was one of the alternative approaches whose studies initiated in the early ‘70. By inducing an adaptive immune response against molecules involved in gamete production, recognition or embryo development, researchers thought that it may be possible to cause infertility. Ideally, an immune response should have been highly specific, possibly long-term but not permanent, and without side effects.

At the moment, promising studies have shown that an anti-GnRH monoclonal antibody inhibits mouse gamete production and blocks fertilization *in vivo*[110]. The same paper discusses the monoclonal antibody IE-3, which reacts to ZP2, able to inhibit fertilization up

to 50% in mice[110]. A novel vaccine against human chorionic gonadotropins (hCG) that impairs embryo development has recently reached Phase II of clinical trials.

Proteins of the egg coat have always been considered optimal candidates for development of immunocontraceptive strategies, because of their tissue specificity and roles during gamete recognition. Passive immunizations with antibodies targeting the ZP indeed showed that it was possible to block fertilization *in vitro*[111] and to reduce animal fertility *in vivo*[112,113]. These findings were further strengthened by experiments based on the cross-reactivity properties of pig ZP that elicited an immune response and can generate antibodies able to bind heterologous ZPs[114]. This finding became a pioneer study for development of ZP vaccines and accordingly, injections of heat-solubilized pig ZP resulted in a significant reduced fertility in rabbits[115], horses[116], dogs[117], primates[118], several ungulate species[119] and elephants[120]. Contraceptive efficacy generally depends on the strength of the individual immune response and on the antibody titers as evidenced by a return of fertility when circulating antibodies decrease.

Together with native pZP, immunizations using recombinant full-length proteins or ZP peptides of ZP1[121,122], ZP2[123] and ZP3[124,125] were shown to reduce fertility *in vivo*.

However, most of the immunizations performed with these formulations (native pig ZP, recombinant ZP protein or ZP peptides) induced severe ovarian dysfunctions and inflammation causing loss of primordial follicles[126].

Although side effects vary across animal species, the outcome discouraged research on ZP epitopes for potential future application in human biology and left dubious questions regarding the efficacy and mechanisms of action of the pig ZP vaccine.

AIMS

"Basically, I have been compelled by curiosity." — Mary Leakey

General aim

The overall aim of this thesis is to investigate the role of the egg coat subunit in gamete recognition with x-ray crystallography.

More specifically, I attempted to:

I Investigate the potential role of sp56 as binding partner of ZP3 and to reconstitute the heterocomplex *in vitro*.

II Help to demonstrate the evolutionary link between invertebrates and mammals showing that egg coat structures share common features by solving the structure of the first ZP-N repeat of ZP2.

III Characterize the VERL-lysin protein complex to elucidate the mechanism of gamete recognition and species-specificity in abalones.

IV Develop a method to express milligrams of protein for structural studies using eukaryotic cells.

IV Elucidate the relation between ZP glycoproteins and future contraceptive strategies and study their impact on female infertility.

METHODOLOGY

This thesis includes several methods spanning from molecular and cell biology, to biochemistry and structural biology. In this section, I briefly described some principles of the methodologies used in this work to better provide an overview for discussions presented in; Study I (detection of protein-protein interaction); Study I-IV (expression of glycoproteins in eukaryotic system); Study III (usage of MBP for initial phasing); Study IV (biophysical characterization by mass spectrometry and *in vitro* fertilization). For more detailed information about methods, please refer to the respective papers.

3.1 Identification of protein-protein interactions

Protein-protein interactions (PPI) play an essential role in biology. Gamete recognition in mammals has been suggested to be an event involving proteins that interact transiently[127,128]. The strength of PPI is measured by the dissociation constant (Kd) which is defined as the ratio between the constant of dissociation reaction (off rate: k_{off}) and the association reaction (on rate: k_{on}) (Kd=k_{on}/k_{off}). When k_{off} is high the reaction proceeds in the reverse direction towards the unbound free forms, so that complex subunits dissociate rapidly. A low k_{on} value associate to weak binding events. Typically, transient interactions present a dissociation constant in the micromolar range and most likely exist *in vivo* as multiple dynamic oligomeric states[129]. On the contrary, strong transient complexes have nanomolar Kd and interaction between the two or more parties has a longer lifetime.

In vitro studies of weak interactions can be technically challenging, because complexes dissociate frequently and present show half-lives (sometimes only for fraction of seconds). To overcome this issue, several studies have shown that weak complexes can be successfully detected by increasing binding avidity[130]. To achieve this, recombinant proteins are engineered and expressed as multiple copies and coupled to beads to increase the locale concentration of binding domains[131] (i.e. Juno-Izuno is 12 μM, or CD2/CD48 90 μM[132]).

In our work we assessed PPI by pull-down assay, Far WB and microscale thermophoresis analysis (MST). The first two methods have in common the ability to detect physical interactions between two or more proteins. Since sp56 is naturally expressed as hexamer, we expected to detect low affinity interactions by pull-down assay due to its high-oligomeric state. Unlike the first two methods, MST measures how molecules move in a temperature gradient, which strongly depends on a variety of molecular properties such as size, charge,

hydration shell or conformation. In this set up, one protein needs to be purified and labelled to a specific fluorophore, while the binding partner is used as a titrant. Based on the thermophoresis path and the concentration of the ligand, it is possible to estimate the Kd of their interaction.

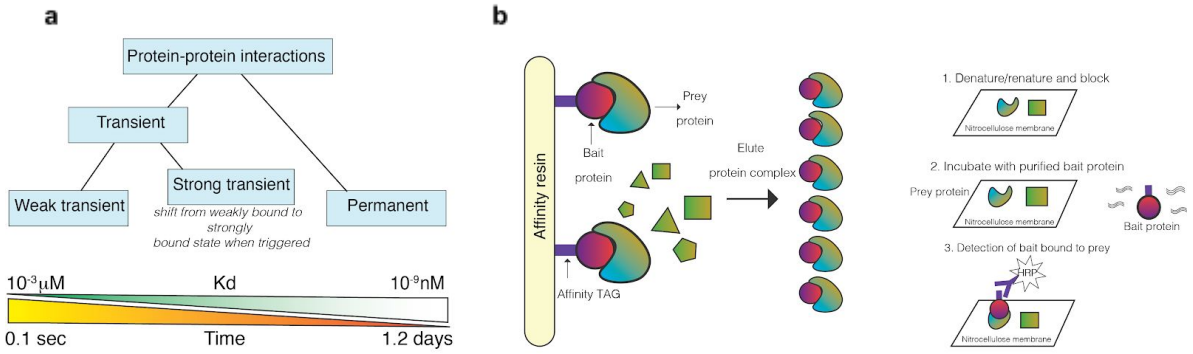

Figure 7. Methods for detection of protein-protein interaction. Panel **a** describes different types on PPI and their relation to Kd and time[133]. b) Graphical description of pull down assay and Far WB.

3.2 Expression of glycoproteins in mammalian and insect cells

Most of the proteins described in this thesis require posttranslational modifications and disulfide bonds for correct folding and/or activity. Glycosylations are defined N-linked when glycans are covalently attached to the nitrogen of asparagine (consensus sequence N-X-S/T) or O-linked when involves the hydroxyl oxygen of a serine or threonine. All N-linked glycans present a common pentasaccharide core made of two *N*-acetylglucosamine (GlcNac) and mannose (Man) ($Man_3GlcNAc_2$) with the most common *N*-glycosidic bond GlcNAcβ1-Asn. Also O-linked glycans contain initial GlcNac residues, but are more heterogeneous.

To ensure the expression of properly folded glycoproteins, we expressed them as secreted proteins by transient transfection in eukaryotic cells (mammalian of insect cells). These two expression systems differ in the type of glycosylation that they produce. Specifically, mammalian cells produce complex type of carbohydrates whose glycan composition might vary according to the cell line used (i.e. HEK293 and CHO-K1). Since glycosylation hinders protein crystallization, we remove the site by either mutating the asparagine residue or by enzymatic digestion using N-glycosidase F (PNGase-F). However, our most successful strategy is to express glycoproteins in GlcNAc transferase I ($GnTI^-$) deficient (HEK293S) cells that exclusively add high-mannose glycans. As a result, glycoproteins become

susceptible to digestion by Endo H, a glycosidase that cleaves in between the first two GlcNAc.

Like HEK293S, insect cells (*Spodoptera frugiperda* and *Trichoplusia ni)* also secrete glycoproteins whose glycans are mostly oligomannose.

3.3 X-ray crystallography: what does the phase problem mean?

Most protein structures have been solved by X-ray crystallography. Its founding principle is Bragg's Law ($2d_{hkl}\sin(\theta)=n\lambda$) of X-ray diffraction by a crystal, a well-ordered packing of molecules in 3D (three-dimensional) space. The building block of a crystal is defined as the unit cell, which is the smallest repeating unit containing the full symmetry. In the protein data bank (PDB), a database containing all structures, only the atomic coordinates of the asymmetric unit are reported. The asymmetric unit and the crystal symmetry contains all the structural information necessary to generated the whole unit cell.

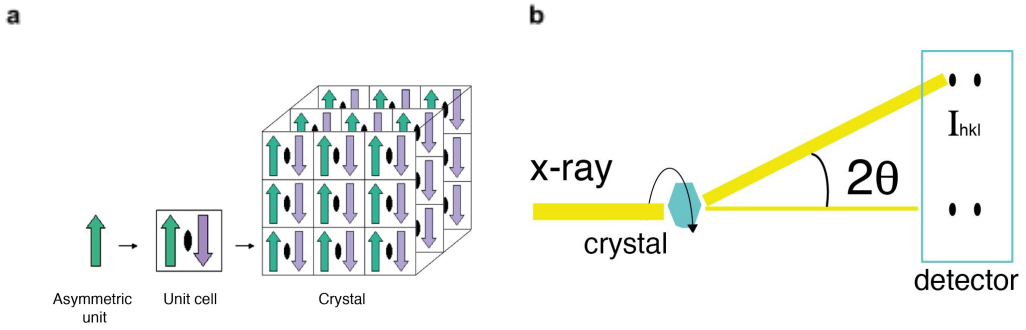

Figure 8. Principles of crystallography. Modified picture taken from PDB. a) Differences between asymmetric unit, unit cells and crystal. b) X-ray scattered by a rotating crystal gives rise to discrete intensities.

When an X-ray beam hits the crystal, the electrons of all the atoms inside the crystal scatter the X-rays in all possible directions. The scattering waves interfere with each other giving rise to a lattice of discrete intensities, whose shape and dimensions are determined by the crystal lattice dimensions and its orientation in relation to the incoming beam and the detector. The intensity of each spot, some are stronger than others, is determined by the unit cell contents, i.e. the protein molecule. The 3D collection of all recorded intensities, which is obtained by rotating the crystal, is called a diffraction pattern and it contains the information of the structural arrangement of all the atoms in the crystal. Each reflection is made of an amplitude (F) and a phase (φ). We can derive the amplitudes from the recorded intensities of the diffraction spots (I), but not their phases. This latter information

is essential to be able to relate the intensities to the position of the atom in the crystal (the phase problem).

The intensities and phases of the crystal are reconstructed, and Fourier transformed to generate an initial electron density map of the crystal. Phases can be inferred from; 1) experimental phasing using heavy atoms which contain a high number of electrons (SIR, MIR, SAD, MAD); 2) computational methods using a related structure as a search model (MR; molecular replacement). The working model is then refined, directed by the electron density map, which continuously improves in iterations until the calculated structure factor amplitudes (Fcalc) come into the closest reasonable agreement with the structure factor amplitudes calculated from the measured intensities (Fobs). In general the appearance and quality of the electron-density map depends on the resolution and quality of the diffraction data and the phases.

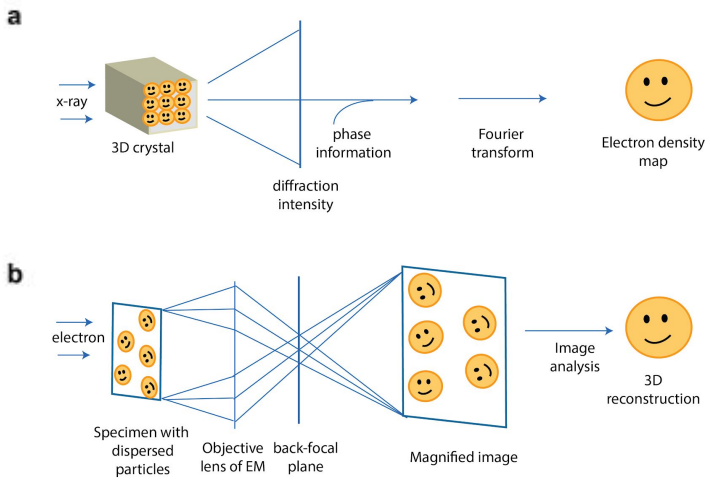

Figure 9. Difference between X-ray crystallography and single particle cryo-electron microscopy (EM)[134] (modified picture from *Wang and Wang* (2017). a) Because of the lack of a focus lens for X-rays, only intensities can be recorded. b) In cryo-EM individual molecules can be imaged directly. In the back-focal plane of the objective lens, the diffraction pattern of a biological sample can also be obtained.

Steps to get well diffracting crystals and determine crystal structure:

1. Engineer a construct and express the protein of interest
2. Get highly purified protein that looks homogeneous and stable
3. Crystallize the protein
4. Collect diffraction data. Poor diffraction data quality can be improved with crystal optimizations (back to point 3). If it fails, crystals benefit from new construct boundaries or improved purity (back to point 1 and 2).

5. Solve the phase problem and calculate the electron density
6. Build a model in the electron density map, refine and validate it

3.4 Matrix assisted laser desorption/ionization - MALDI

In mass spectrometry (MS) sample molecules (typically protein fragments) are co-crystallized with a matrix, the proton source, and excited by a laser light. Upon irradiation, all charged molecules are accelerated by a magnetic field and travel through a mass analyzer to the detector. Since they all receive the same amount of energy, the difference in time-of-flight (TOF) is correlated to the mass (m/z) of the ion. Thus by MS, we get molecular weight information and as a consequence can assess whether post-translational modification are present. However, to obtain the exact sequence composition of a certain fragment it is possible to select a specific ion and fragment it (MS/MS).

3.5 Superovulation in mice and *in vitro* fertilization

For IVF assays 3- to 5-week old C57BL6/J females mice were used as oocyte donors. A standard protocol consist of stimulating females with a intraperitoneal administration of 5 IU Pregnant Mare Serum Gonadotrophin (PMSG; Folligon, Intervet) and, 48 h later, with 5 IU human Chorionic Gonadotropin (hCG; Chorulon, Intervet). 12 h after the hCG injection females were sacrificed and ampullas from individual animals were retrieved and equilibrated in HTF (human tubal fluid) medium. The IVF success rate, number of embryos estimated for this strain is around 65%[135,136].

RESULTS AND DISCUSSION

4.1 Study I

Molecular basis of egg sperm interaction:characterization of hamster sp56

Sp56 was previously identified as the sperm protein possessing ZP3 binding activity in mouse[66]. It belongs to the C3/C4 superfamily of proteins and contains a repetition of CCP domain[15]. Capacitation triggers local fusion events within the acrosomal membranes that lead to increasing exposure of sp56 molecules to the external environment. Accordingly, only capacitated sperm is able to bind and penetrate the ZP[14].

Considering the role of glycosylation in sperm adhesion, we opted to use hamster as a model system and express glycoproteins in CHO-K1 cells. Hamster ZP3 was also previously shown to be recombinantly expressed in this cell line as an active protein, with the ability to induce acrosome reaction of sperm *in vitro*[137].
In this work, we cloned and amplified the homolog of sp56 from hamster testis and expressed it recombinantly first in mammalian cells, and subsequently in insect to attempt to increase the protein yield. In both cases, sp56 appeared to be secreted as a properly folded protein because it formed a high molecular weight oligomer of 250 kDa, as expected for a member of the C3/C4 protein family. Recombinant hamster ZP3 was also expressed, and its interaction with sp56 was assessed by pull-down assays and Far WB.
However, our biochemical approach failed to detect a convincingly hetero-complex sp56-ZP3. This might suggest that ZP3-sp56 interaction is transient and labile or might involve additional multiple partners in hamster that haven't been identified yet.

Although homozygous male null mice for sp56 reported during the course of this Ph.D. studies were found fertile[138], several additional analyses might further be considered to establish whether the absence of sp56 influences fertilization at the level of sperm recognition; for instance by using a lower sperm number and possibly apply competition pressure, or by testing the acrosome reaction using native ZP as a more physiological inducer. Advances in genome editing manipulation like Crisp/Cas9 might also be an additional tool to establish the *in vivo* role of sp56 in other animal species.

4.2 Study II

Structural basis of egg coat-sperm recognition at fertilization

It was previously suggested that sperm-binding protein VERL and ZP2 are protein homologs[83]. By applying bioinformatics analysis, both structures were predicted to contain a repetition of ZP-N domain in their N-termini. However, considering their insignificant sequence identity (~10 %), the hypothesis remained highly speculative.

In Study II we experimentally confirmed this hypothesis and proved that both egg coat sperm receptors in invertebrates (abalone VERL) and mammals (mouse ZP2) share a common ZP-N domain to accomplish essential functions during fertilization. Although they fold similarly, the structures present some differences, most likely as a consequence of the type of the fertilization mechanisms.

Unlike the first repeat, VERL repeat two and three bind to lysin as confirmed by MST and gel filtration analysis. Crystallographic analysis of VERL-lysin complexes showed VERL and lysin engaged in a 1:1 complex held by an extensive amphipathic interface. The binding mostly involves the β-strands B, D-E and the loop ee' of VERL. Interestingly, different from what observed in the VERL structure, the β-strand D of ZP2-N1 locates in the opposite β-sheet, close to β-strand E. Since no binding partner for ZP2 has been identified yet, we speculated that either the protein evolved a different way to bind sperm, or that ZP2 lost the binding site and therefore it indirectly regulates sperm binding by changing the structure of the ZP.

VE dissolution assays performed with lysin from different species was shown to be species-specific[78] (see introduction 1.8). By performing pull-down analysis with red, pink and black lysin, we could identified VR2 as the key component that regulates species-specificity. VR2 binds lysin with lower affinity compared to VR3 and mutation in N and C-termini of lysin affected binding exclusively to VR2.

By analyzing and comparing the crystal packing of the unbound and bound VERL structures, we could derived a model of how filaments are organized within the egg coat and proposed a mechanism of how lysin makes a hole in a non-enzymatic manner. Interestingly, our proposed mechanism recapitulates the VE dissolution assays in which lysin slowly opens up the filament by interacting with VR2 first. When VR3 binding sites become accessible, the sperm protein titrate all the remaining repeats by disrupting the egg coat structure.

4.3 Study III

Easy mammalian expression and crystallography of maltose-binding protein-fused human proteins

Glycoproteins require post-translational modifications to correctly fold and are thus often expressed in mammalian or insect cells. However protein expression using these types of cells is often not sufficient for conducting structural studies.

Study III describes the advantage of using bacterial maltose-binding protein (MBP) as a fusion partner to express considerable amounts of glycoproteins in mammalian cells. MBP was codon-optimized for better expression in human cells (mMBP). Based on published results, a series of amino acid substitutions were introduced in mMBP to increase its solubility, maltose binding affinity and ability to crystallize. The fusion to mMBP enhanced expression and stability of five challenging glycoproteins that could not being otherwise exploited for structural studies. Most importantly, mMBP provides phasing information for solving crystal structures by molecular replacement. This approach led to the structure determination of five glycoproteins (UMOD, VR1, VR3 and two constructs of ENG) where mMBP was used as a successful search model. Analysis of their crystal packing reveals that most of the interactions are mediated by mMBP loops and N-terminal tag. Interestingly, direct comparison of UMOD fused to mMBP or human growth hormone (hGH), an alternative system suggested to increase protein expression, showed that mMBP substantially increased UMOD expression (Figure 9). Introduction of a specific cleavage site, in our case TEV protease (tobacco etch virus), would ensure removal of mMBP. More interesting, considering that one selenomethionine for every 75-100 amino acids[139] is enough for conducting successful derivative experiment, the content of methionine residues in mMBP (6 for a 380 residues protein) might offer the possibility to exploit usage of selenomethionine derivatives[140] for structure determination of larger passenger protein.

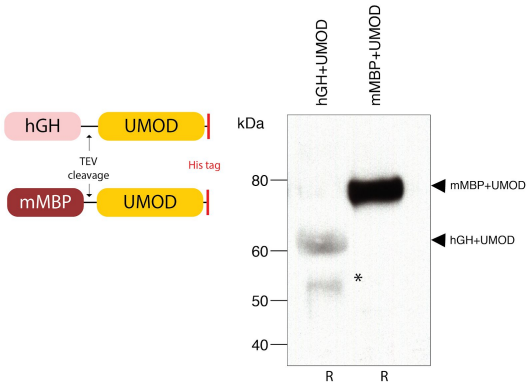

Figure 9. Small scale expression of UMOD fused to mMBP or hGH in HEK293T. 20 µl of medium containing hGH-UMOD and MBP-UMOD were analyzed in reducing condition by immunoblotting with monoclonal anti-His. Notably, non-specific cleavage occurred in hGH-UMOD (asterisk).

4.4 Study IV

Molecular insights into ZP-based contraception and ZP-linked infertility

Contraception and female infertility are two specular aspects of reproduction that involve the ZP. Understanding how egg coat subunits functions and assemble would certainly have an impact in the field of reproductive medicine and help to clarify the molecular mechanisms behind these two events.

To address these points we first determined the crystal structure of a ZP-antigen in complex with a contraceptive monoclonal antibody. Using biochemical assays together with biophysical characterizations, we highlighted the importance of maintaining the 3D conformation of the epitope to ensure high binding affinity of the antibody. Based on the crystal structure, we could engineer a construct encompassing the regions of the antibody involved in antigen binding and test it in IVF. Parallel studies of ZP proteins carrying mutations found in infertile patients revealed critical features that interfere with structure of the egg coat. Targeting these elements might lead to a new generation of contraceptive agents.

CONCLUSION AND FUTURE PERSPECTIVE

The egg coat matters

In this thesis, several glycoproteins involved in fertilization have been studied. We started by looking at fertilization in mammals. Data supporting the role of sp56 as a binding partner for ZP3 led us to investigate more in details their putative interaction. Considering the natural high avidity of sp56, I expected to detect even low-binding affinity events, as in the case of previous weak complex characterizations[141]. The inability to identify a hetero-complex, together with the finding by another research group that sp56 null mice are fertile[138], suggested that the egg coat-sperm binding is either redundant. Nonetheless, additional experiments, as described in the section of Study I, may further highlight the function of sp56 in mouse fertilization *in vivo*.

In this research, it often occurred that *in vivo* studies did not recapitulate the findings obtained *in vitro*, mostly assessed by working with individual purified proteins. This discrepancy may be explained by the fact that results have mainly been biased by mouse models. This might suggest that investigating and comparing fertilization mechanisms from several species may identify novel features and help to understand *in vivo* function of novel proteins. In light of this, advances in gene-editing manipulations may encourage this type of approach.

Unlike mammals, proteins involved in egg-sperm binding in abalone were already known. By determining crystal structures of VERL-lysin complex, we explained at the atomic level events prior to sperm penetration. We also provided a definitive link between invertebrate and vertebrate fertilization by showing that egg coat proteins adopt similar 3D structures. These findings had significant implications for mammalian fertilization. The structural differences between ZP2 and VERL in the area corresponding to the binding site for lysin would argue against the existence of such a molecule in mammalian sperm. By releasing abundant amount of lysin, sperm makes a hole and directly goes through the egg coat to fertilize the egg and accordingly, VE dissolution in abalone completes in 20 minutes, while sperm penetration in mammals takes roughly 4 hours. In mammals, a lysin-like molecule might have been lost during evolution, possibly as a result of the internal fertilization or the increased thickness of the ZP (7 µm-24 µm versus 0.6 µm in abalone). Considering the increased thickness of the matrix, a solely non-enzymatic mechanism would not be optimal for sperm to penetrate the ZP. As recently suggested, mammalian sperm penetration

combines both mechanical forces[17] and an enzymatic activity[13] in which acrosin, a serine protease abundant in the acrosome, has been suggested to play a significant role[142].

Identification of proteins with ZP-binding activity might clarify whether sperm really interacts with the mammalian ZP or just passes through by enzymatically cleaving ZP filaments. Considering data suggesting that the ZP structure modulates sperm binding, it would be ideal to perform biochemical studies using purified egg coat filaments containing all ZP subunits rather than individual subunits. Alternatively, a characterization of spermatozoa bound to the ZP might help to identify which features are possibly involved. For this purpose, binding-competent sperm cells could be selected either by using the standard sperm-binding assay employed to test the function of ZP3[143] or using oocytes expressing human ZP2[60].

Crystal structures of egg coat glycoproteins was achieved by fusing them to mMBP. In Study III we described the advantages of such method to express large amounts of glycoproteins using mammalian cells. So far, 65 different users have requested the pHLmMBP vectors from Addgene, and a recently reported study has shown that fusion to mMBP helped the expression of a peptide[144]. To release mMBP and obtain untagged glycoproteins, we described the usage of a TEV cleavage site as an option. This site can be replaced by other sequences depending on the experimental needs and preferences (see a full list here; List of proteases[145]).

Finally, considering the important role of ZP in human infertility, we attempted to characterize some of the mutations affecting ZP genes. Genotyping more infertile patients and increasing knowledge of the regulation and expression of ZP subunits in mammalian oocytes might provide better insights into how ZP matrix is assembled.

ACKNOWLEDGEMENTS

This project owes so much to many creative and beautiful people who encouraged and inspired me over these years. For this, I would like to sincerely express my gratitude to my supervisor Prof. **Luca Jovine** who accepted me as a student and introduced me to the world of ZP proteins. Your knowledge, perfectionism, and passion for science are truly inspiring, and I am sure that will help me to succeed in the future. I learned how to be critical, open-minded and, how to also design experiments with at least 4-5 controls =). Thanks!

Thanks to all the **LJO** group, the past and the present, who made my journey so incredible and a great learning experience. It has been full of ups and downs, and you have always been there to cheer me up. In particular, thanks to; **Ling Han** who shared the secrets for getting perfect western blots; to **Hamed Sadat** - as he said it once "my partner in crime." I have missed your nagging and stories about gym exercises in these last two years =); to **Marcel Bokhove** for having shared his eclectic passion for music and knowledge of crystallography; to **Amy Cherry** for having taught perseverance and English (hopefully); to **Takako Saito** for showing me how to keep calm and positive; to **Isha Raj** for being my indian feminist - thanks you so much for the help and support during all my writings; to **Romina Croci** for the nice italian chats; **Kaoru Nishimura** for having shared your expensive saké; to **Blanca Algarra** and **Sara Zamora** for the spanish language lesson and fun in the lab; to **Dirk Fahrenkamp** who took over the ZP2 project =) - thanks for the support, the laugh and fun moments; to **Eileen Dietzel** who helped me a lot in this last year, for being always available to answer to all my questions about reproduction and the fabulous cakes.

Thanks to Prof. **Tsukasa Matsuda, Shunsuke Nishio** and **Daniele de Sanctis** for having collaborated on the projects of this thesis and for sharing your knowledge and expertise.

I would like to acknowledge **José Inzunza**, **Mona Farthoo**, **Michael Zmarzlak**, **Daniel Bergström**, **Maria Kasper**, **Alexandra Are** and **Stephan Teglund** for help with animal experiments and IVFs.

Thanks to my dear friends that I have made over these years in Stockholm! Thanks to **Alessandra Villa** for being the most inspiring woman in science that I know and for all the precious advices. Thanks to my big brothers **Giuseppe Baldi** and **Pablo Hernández Varas** for their love, their never-ending conversations and the support; to my dearest

friends that made me feel home since day one and fill my life with happiness; grazie a **Erica Miraglia**, **Lorenza Capantini**, **Simona Conte**, **Valentina Giai**, **Valentina Pampanini**, **Alice Ghidini**, **Carolina Forte Rigos and Mustafa Ismail.** Thanks to **Valentino Davio** for his insights and opinions and for helping me together with **Cristian Mariolini** to organize a memorable post Ph.D celebration. To **Rajesh Ponnusamy** and **Dominick Possner** for their support and conversations about structural biology. Thank you **Ana Amaral** for your insights and help while I was writing the thesis. To **Andreea Scacioc** and **Tales Rocha de Moura** with who I share my "nerdiness" and the last sensational piece of science published, thank you. A big thanks goes to my climbing partner **Sylvie le Guyader** with whom I have spent lot of beautiful time - maybe a bit less in the last two years.

Thanks to my sweet friends who live in Italy or around the world to inspire me and having shared my long path in science: **Letizia Assinnato**, **Alessia Bagnobianchi**, **Mariagrazia Caccamo**, **Sara Chiaretti**, **Alessia Braione** and **Federica Cossu**. Thanks to the fantastic group of medicinal chemistry of NMS - **Teresa, Stefano**, **Fabio**, **Dannica**, **Federico**, **Giovanni** for having shared your knowledge and fun at work and to **Elena Casale**, **Roberto Bossi** and **Jay Bertrand** for believing in me; to **Eva Pistarino** for helping me to find a place when I decided to move to Stockholm and to **Christian Orrenius** for his tips of Stockholm.

Thanks to **Fatou Hellman** and **Morgan Berglund** to be part of my life and your interest to my work :).

Thanks to the whole **Kjellgren family** who welcomed me and support me all the time and make me feel very fortunate.

Un grazie speciale va all'intera mia famiglia, **Ciro**, **Melina**, **Marco**, **Davide**, **Caterina**, **Giulia**, **Mirko, Pietro, Grazia, Luca, Giulia, Angela, Natale** il cui supporto e ammirazione riesce a percorrere kilometri e kilometri di distanza. Ai miei due fratelli Davide e Marco dedico questo mio traguardo.

As last but not least, a very big thank to my beloved **Alex Kjellgren** who together with **Enzino** are the most important people of my life. You have listened all my stories about ZP and sperm-binding proteins. This thesis would have never been possible without you!

REFERENCES

1. Wong, J. L. & Wessel, G. M. Defending the zygote: search for the ancestral animal block to polyspermy. *Curr. Top. Dev. Biol.* **72,** 1–151 (2006).
2. Carroll, E. J., Wei, S. H., Nagel, G. M. & Ruibal, R. Structure and macromolecular composition of the egg and embryo jelly coats of the anuran Lepidobatrachus laevis. *Dev. Growth Differ.* **33,** 37–43 (1991).
3. Wassarman, P. M., Liu, C. & Litscher, E. S. Building the mammalian zona pellucida: an organelle that regulates fertilization. *Microscopy of Reproduction and Development: a dynamic approach (Edited by: Motta PM) Rome: Antonio Delfino Editore* 169–176 (1997).
4. Izquierdo-Rico, M. J., Gimeno, L., Jiménez-Cervantes, C., Ballesta, J. & Avilés, M. Biosynthesis of hamster zona pellucida is restricted to the oocyte. *Theriogenology* **75,** 463–472 (2011).
5. Gook, D. A., Edgar, D. H., Borg, J. & Martic, M. Detection of zona pellucida proteins during human folliculogenesis. *Hum. Reprod.* **23,** 394–402 (2008).
6. Lee, V. H. & Dunbar, B. S. Developmental expression of the rabbit 55-kDa zona pellucida protein and messenger RNA in ovarian follicles. *Dev. Biol.* **155,** 371–382 (1993).
7. Sinowatz, F., Kölle, S. & Töpfer-Petersen, E. Biosynthesis and expression of zona pellucida glycoproteins in mammals. *Cells Tissues Organs* **168,** 24–35 (2001).
8. Nishio, S. & Matsuda, T. Fertilization 1: Sperm-Egg Interaction. *Adv. Exp. Med. Biol.* **1001,** 91–103 (2017).
9. Sasanami, T. *Avian Reproduction: From Behavior to Molecules*. (Springer, 2017).
10. Gadella, B. M. & Visconti, P. E. Regulation of capacitation. in *The Sperm Cell* (eds. De Jonge, C. J., Barratt, C., De Jonge, C. J. & Barratt, C.) 134–169 (Cambridge University Press, 2006).
11. Sutton, K. A., Jungnickel, M. K. & Florman, H. M. A polycystin-1 controls postcopulatory reproductive selection in mice. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 8661–8666 (2008).
12. Sumigama, S. *et al.* Progesterone Accelerates the Completion of Sperm Capacitation and Activates CatSper Channel in Spermatozoa from the Rhesus Macaque. *Biol. Reprod.* **93,** 130 (2015).
13. Honda, A., Siruntawineti, J. & Baba, T. Role of acrosomal matrix proteases in sperm–zona pellucida interactions. *Hum. Reprod. Update* **8,** 405–412 (2002).
14. Buffone, M. G., Rodriguez-Miranda, E., Storey, B. T. & Gerton, G. L. Acrosomal exocytosis of mouse sperm progresses in a consistent direction in response to zona pellucida. *J. Cell. Physiol.* **220,** 611–620 (2009).
15. Wassarman, P. M. Mammalian fertilization: the strange case of sperm protein 56. *Bioessays* **31,** 153–158 (2009).
16. Flesch, F. M. & Gadella, B. M. Dynamics of the mammalian sperm plasma membrane in the process of fertilization. *Biochim. Biophys. Acta* **1469,** 197–235 (2000).
17. Miyata, H. *et al.* Sperm calcineurin inhibition prevents mouse fertility with implications for male contraceptive. *Science* **350,** 442–445 (2015).
18. Bianchi, E., Doe, B., Goulding, D. & Wright, G. J. Juno is the egg Izumo receptor and is essential for mammalian fertilization. *Nature* **508,** 483–487 (2014).
19. Inoue, N., Ikawa, M., Isotani, A. & Okabe, M. The immunoglobulin superfamily protein Izumo is required for sperm to fuse with eggs. *Nature* **434,** 234–238 (2005).
20. Tosti, E. & Boni, R. Electrical events during gamete maturation and fertilization in animals and humans. *Hum. Reprod. Update* **10,** 53–65 (2004).
21. Que, E. L. *et al.* Zinc sparks induce physiochemical changes in the egg zona pellucida that prevent polyspermy. *Integr. Biol.* **9,** 135–144 (2017).
22. Liu, M. The biology and dynamics of mammalian cortical granules. *Reprod. Biol. Endocrinol.* **9,** 149 (2011).
23. Burkart, A. D., Xiong, B., Baibakov, B., Jiménez-Movilla, M. & Dean, J. Ovastacin, a cortical granule protease, cleaves ZP2 in the zona pellucida to prevent polyspermy. *J. Cell Biol.* **197,** 37–44 (2012).
24. Gahlay, G., Gauthier, L., Baibakov, B., Epifano, O. & Dean, J. Gamete recognition in mice depends on the cleavage status of an egg's zona pellucida protein. *Science* **329,** 216–219 (2010).
25. Bleil, J. D. & Wassarman, P. M. Synthesis of zona pellucida proteins by denuded and follicle-enclosed mouse oocytes during culture in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **77,** 1029–1033 (1980).
26. Chapman, J. A., Leigh, C. M. & Breed, W. G. The zona pellucida of the koala (Phascolarctos cinereus): its morphogenesis and thickness. *J. Anat.* **209,** 393–400 (2006).
27. Eppig, J. J. Reproduction: Oocytes Call, Granulosa Cells Connect. *Curr. Biol.* **28,** R354–R356 (2018).
28. Jovine, L., Qi, H., Williams, Z., Litscher, E. S. & Wassarman, P. M. A duplicated motif controls assembly of zona pellucida domain proteins. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 5922–5927 (2004).
29. Jovine, L., Darie, C. C., Litscher, E. S. & Wassarman, P. M. Zona pellucida domain proteins. *Annu. Rev. Biochem.* **74,** 83–114 (2005).
30. Bokhove, M. & Jovine, L. Structure of Zona Pellucida Module Proteins. in *Current Topics in Developmental Biology* (Academic Press, 2018).
31. Rankin, T. L. *et al.* Defective zonae pellucidae in Zp2-null mice disrupt folliculogenesis, fertility and development. *Development* **128,** 1119–1126 (2001).
32. Rankin, T. *et al.* Mice homozygous for an insertional mutation in the Zp3 gene lack a zona pellucida and are infertile. *Development* **122,** 2903–2910 (1996).
33. Bleil, J. D., Greve, J. M. & Wassarman, P. M. Identification of a secondary sperm receptor in the mouse

egg zona pellucida: role in maintenance of binding of acrosome-reacted sperm to eggs. *Dev. Biol.* **128,** 376–385 (1988).
34. Baibakov, B., Boggs, N. A., Yauger, B., Baibakov, G. & Dean, J. Human sperm bind to the N-terminal domain of ZP2 in humanized zonae pellucidae in transgenic mice. *J. Cell Biol.* **197,** 897–905 (2012).
35. Monné, M., Han, L., Schwend, T., Burendahl, S. & Jovine, L. Crystal structure of the ZP-N domain of ZP3 reveals the core fold of animal egg coats. *Nature* **456,** 653–657 (2008).
36. Monné, M. & Jovine, L. A structural view of egg coat architecture and function in fertilization. *Biol. Reprod.* **85,** 661–669 (2011).
37. Han, L. *et al.* Insights into egg coat assembly and egg-sperm interaction from the X-ray structure of full-length ZP3. *Cell* **143,** 404–415 (2010).
38. Jovine, L., Qi, H., Williams, Z., Litscher, E. & Wassarman, P. M. The ZP domain is a conserved module for polymerization of extracellular proteins. *Nat. Cell Biol.* **4,** 457–461 (2002).
39. Bokhove, M. *et al.* A structured interdomain linker directs self-polymerization of human uromodulin. *Proc. Natl. Acad. Sci. U. S. A.* **113,** 1552–1557 (2016).
40. Conner, S. J. & Hughes, D. C. Analysis of fish ZP1/ZPB homologous genes--evidence for both genome duplication and species-specific amplification models of evolution. *Reproduction* **126,** 347–352 (2003).
41. Moros-Nicolás, C. *et al.* Analysis of ZP1 gene reveals differences in zona pellucida composition in carnivores. *Reprod. Fertil. Dev.* (2017). doi:10.1071/RD17022
42. Greve, J. M. & Wassarman, P. M. Mouse egg extracellular coat is a matrix of interconnected filaments possessing a structural repeat. *J. Mol. Biol.* **181,** 253–264 (1985).
43. Rankin, T., Talbot, P., Lee, E. & Dean, J. Abnormal zonae pellucidae in mice lacking ZP1 result in early embryonic loss. *Development* **126,** 3847–3855 (1999).
44. Lefievre, L. *et al.* Identification of four glycoproteins in the human zona pellucida by proteomic analysis of five oocytes. *Hum. Reprod.* **19,** 1580–1586 (2004).
45. Yauger, B., Boggs, N. A. & Dean, J. Human ZP4 is not sufficient for taxon-specific sperm recognition of the zona pellucida in transgenic mice. *Reproduction* **141,** 313–319 (2011).
46. Avella, M. A., Baibakov, B. & Dean, J. A single domain of the ZP2 zona pellucida protein mediates gamete recognition in mice and humans. *J. Cell Biol.* **205,** 801–809 (2014).
47. von Witzendorff, D. *et al.* Characterization of the acidic N-linked glycans of the zona pellucida of prepuberal pigs by a mass spectrometric approach. *Carbohydr. Res.* **344,** 1541–1549 (2009).
48. Bleil, J. D. & Wassarman, P. M. Mammalian sperm-egg interaction: identification of a glycoprotein in mouse egg zonae pellucidae possessing receptor activity for sperm. *Cell* **20,** 873–882 (1980).
49. Florman, H. M., Bechtol, K. B. & Wassarman, P. M. Enzymatic dissection of the functions of the mouse egg's receptor for sperm. *Dev. Biol.* **106,** 243–255 (1984).
50. Bleil, J. D. & Wassarman, P. M. Autoradiographic visualization of the mouse egg's sperm receptor bound to sperm. *J. Cell Biol.* **102,** 1363–1371 (1986).
51. Vazquez, M. H., Phillips, D. M. & Wassarman, P. M. Interaction of mouse sperm with purified sperm receptors covalently linked to silica beads. *J. Cell Sci.* **92 (Pt 4),** 713–722 (1989).
52. Chinoy, Z. S., Friscourt, F., Capicciotti, C. J., Chiu, P. & Boons, G.-J. Chemoenzymatic Synthesis of Asymmetrical Multi-Antennary N-Glycans to Dissect Glycan-Mediated Interactions between Human Sperm and Oocytes. *Chem.--Eur. J.* (2018).
53. Florman, H. M. & Storey, B. T. Mouse gamete interactions: the zona pellucida is the site of the acrosome reaction leading to fertilization in vitro. *Dev. Biol.* **91,** 121–130 (1982).
54. Gong, X., Dubois, D. H., Miller, D. J. & Shur, B. D. Activation of a G protein complex by aggregation of beta-1, 4-galactosyltransferase on the surface of sperm. *Science* (1995).
55. Arnoult, C., Cardullo, R. A., Lemos, J. R. & Florman, H. M. Activation of mouse sperm T-type Ca2+ channels by adhesion to the egg zona pellucida. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 13004–13009 (1996).
56. Moller, C. C. & Wassarman, P. M. Characterization of a proteinase that cleaves zona pellucida glycoprotein ZP2 following activation of mouse eggs. *Dev. Biol.* **132,** 103–112 (1989).
57. Tian, J., Gong, H. & Lennarz, W. J. Xenopus laevis sperm receptor gp69/64 glycoprotein is a homolog of the mammalian sperm receptor ZP2. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 829–834 (1999).
58. Rankin, T. L. *et al.* Fertility and taxon-specific sperm binding persist after replacement of mouse sperm receptors with human homologs. *Dev. Cell* **5,** 33–43 (2003).
59. Dietzel, E., Floehr, J. & Jahnen-Dechent, W. The biological role of fetuin-B in female reproduction. *Ann Reprod Med Treat* **1,** 1003 (2016).
60. Baibakov, B., Gauthier, L., Talbot, P., Rankin, T. L. & Dean, J. Sperm binding to the zona pellucida is not sufficient to induce acrosome exocytosis. *Development* **134,** 933–943 (2007).
61. Avella, M., Baibakov, B., Burkart Sadusky, A. & Dean, J. ZP2 peptide beads that select human sperm in vitro, decoy mouse sperm in vivo, and provide reversible contraception. *Fertil. Steril.* **106,** e112 (2016).
62. Visconti, P. E. & Florman, H. M. Mechanisms of sperm-egg interactions: between sugars and broken bonds. *Sci. Signal.* **3,** e35 (2010).
63. Wassarman, P. M. & Litscher, E. S. Mammalian fertilization: the egg's multifunctional zona pellucida. *Int. J. Dev. Neurosci.* (2004).
64. Bedford, J. M. Sperm/egg interaction: the specificity of human spermatozoa. *Anat. Rec.* **188,** 477–487 (1977).

65. Snell, W. J. & White, J. M. The molecules of mammalian fertilization. *Cell* **85,** 629–637 (1996).
66. Bleil, J. D. & Wassarman, P. M. Identification of a ZP3-binding protein on acrosome-intact mouse sperm by photoaffinity crosslinking. *Proc. Natl. Acad. Sci. U. S. A.* **87,** 5563–5567 (1990).
67. Buffone, M. G., Foster, J. A. & Gerton, G. L. The role of the acrosomal matrix in fertilization. *Int. J. Dev. Biol.* **52,** 511–522 (2008).
68. Cheng, A. *et al.* Sperm-egg recognition in the mouse: characterization of sp56, a sperm protein having specific affinity for ZP3. *J. Cell Biol.* **125,** 867–878 (1994).
69. Bookbinder, L. H., Cheng, A. & Bleil, J. D. Tissue-and species-specific expression of sp56, a mouse spern fertilization protein. *Science* **269,** 86 (1995).
70. Bleil, J. D. & Wassarman, P. M. Galactose at the nonreducing terminus of O-linked oligosaccharides of mouse egg zona pellucida glycoprotein ZP3 is essential for the glycoprotein's sperm receptor …. *Proceedings of the National* (1988).
71. Hirohashi, N., Gerton, G. L. & Buffone, M. G. Video imaging of the sperm acrosome reaction during in vitro fertilization. *Commun. Integr. Biol.* **4,** 471–476 (2011).
72. Hofmeyer, T. *et al.* Arranged sevenfold: structural insights into the C-terminal oligomerization domain of human C4b-binding protein. *J. Mol. Biol.* **425,** 1302–1317 (2013).
73. Buffone, M. G. *et al.* Recombinant mouse sperm ZP3-binding protein (ZP3R/sp56) forms a high order oligomer that binds eggs and inhibits mouse fertilization in vitro. *J. Biol. Chem.* **283,** 12438–12445 (2008).
74. Buffone, M. G., Kim, K.-S., Doak, B. J., Rodriguez-Miranda, E. & Gerton, G. L. Functional consequences of cleavage, dissociation and exocytotic release of ZP3R, a C4BP-related protein, from the mouse sperm acrosomal matrix. *J. Cell Sci.* **122,** 3153–3160 (2009).
75. Lewis, C. A., Talbot, C. F. & Vacquier, V. D. A protein from abalone sperm dissolves the egg vitelline layer by a nonenzymatic mechanism. *Dev. Biol.* **92,** 227–239 (1982).
76. Leighton, D. L. & Lewis, C. A. Experimental hybridization in abalones. *International Journal of Invertebrate Reproduction* **5,** 273–282 (1982).
77. Swanson, W. J. & Vacquier, V. D. The abalone egg vitelline envelope receptor for sperm lysin is a giant multivalent molecule. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 6724–6729 (1997).
78. Lyon, J. D. & Vacquier, V. D. Interspecies chimeric sperm lysins identify regions mediating species-specific recognition of the abalone egg vitelline envelope. *Dev. Biol.* **214,** 151–159 (1999).
79. Metz, E. C., Robles-Sikisaka, R. & Vacquier, V. D. Nonsynonymous substitution in abalone sperm fertilization genes exceeds substitution in introns and mitochondrial DNA. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 10676–10681 (1998).
80. Galindo, B. E., Vacquier, V. D. & Swanson, W. J. Positive selection in the egg receptor for abalone sperm lysin. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 4639–4643 (2003).
81. Shaw, A., McRee, D. E., Vacquier, V. D. & Stout, C. D. The crystal structure of lysin, a fertilization protein. *Science* **262,** 1864–1867 (1993).
82. Aagaard, J. E., Vacquier, V. D., MacCoss, M. J. & Swanson, W. J. ZP domain proteins in the abalone egg coat include a paralog of VERL under positive selection that binds lysin and 18-kDa sperm proteins. *Mol. Biol. Evol.* **27,** 193–203 (2010).
83. Swanson, W. J. *et al.* The molecular basis of sex: linking yeast to human. *Mol. Biol. Evol.* **28,** 1963–1966 (2011).
84. Diller, T. C., Shaw, A., Stura, E. A., Vacquier, V. D. & Stout, C. D. Acid pH crystallization of the basic protein lysin from the spermatozoa of red abalone (Haliotis rufescens). *Acta Crystallogr. D Biol. Crystallogr.* **50,** 620–626 (1994).
85. Shaw, A., Fortes, P. A., Stout, C. D. & Vacquier, V. D. Crystal structure and subunit dynamics of the abalone sperm lysin dimer: egg envelopes dissociate dimers, the monomer is the active species. *J. Cell Biol.* **130,** 1117–1125 (1995).
86. Richardson, S. J., Senikas, V. & Nelson, J. F. Follicular depletion during the menopausal transition: evidence for accelerated loss and ultimate exhaustion. *J. Clin. Endocrinol. Metab.* **65,** 1231–1237 (1987).
87. Hadek, R. The Structure of the Mammalian Egg. in *International Review of Cytology* (eds. Bourne, G. H. & Danielli, J. F.) **18,** 29–71 (Academic Press, 1965).
88. Qi, H., Williams, Z. & Wassarman, P. M. Secretion and assembly of zona pellucida glycoproteins by growing mouse oocytes microinjected with epitope-tagged cDNAs for mZP2 and mZP3. *Mol. Biol. Cell* **13,** 530–541 (2002).
89. Epifano, O., Liang, L. F., Familari, M., Moos, M. C., Jr & Dean, J. Coordinate expression of the three zona pellucida genes during mouse oogenesis. *Development* **121,** 1947–1956 (1995).
90. Gabrielsen, A., Bhatnager, P. R., Petersen, K. & Lindenberg, S. Influence of zona pellucida thickness of human embryos on clinical pregnancy outcome following in vitro fertilization treatment. *J. Assist. Reprod. Genet.* **17,** 323–328 (2000).
91. Shiloh, H. *et al.* The impact of cigarette smoking on zona pellucida thickness of oocytes and embryos prior to transfer into the uterine cavity. *Hum. Reprod.* **19,** 157–159 (2004).
92. Palmstierna, M., Murkes, D., Csemiczky, G., Andersson, O. & Wramsby, H. Zona pellucida thickness variation and occurrence of visible mononucleated blastomeres in preembryos are associated with a high pregnancy rate in IVF treatment. *J. Assist. Reprod. Genet.* **15,** 70–75 (1998).
93. Bertrand, E., Van Den Bergh, M. & Englert, Y. Fertilization and early embryology: Does zona pellucida thickness influence the fertilization rate? *Hum. Reprod.* **10,** 1189–1193 (1995).

94. Shen, Y., Stalf, T., Mehnert, C., Eichenlaub-Ritter, U. & Tinneberg, H.-R. High magnitude of light retardation by the zona pellucida is associated with conception cycles. *Hum. Reprod.* **20,** 1596–1606 (2005).
95. Montag, M. *et al.* Oocyte zona birefringence intensity is associated with embryonic implantation potential in ICSI cycles. *Reprod. Biomed. Online* **16,** 239–244 (2008).
96. Rienzi, L., Vajta, G. & Ubaldi, F. Predictive value of oocyte morphology in human IVF: a systematic review of the literature. *Hum. Reprod. Update* **17,** 34–45 (2011).
97. Ya, R., Que, E. L., O'Halloran, T. V. & Woodruff, T. K. Zinc as a Key Meiotic Cell-Cycle Regulator in the Mammalian Oocyte. in *Zinc Signals in Cellular Functions and Disorders* 315–333 (Springer, Tokyo, 2014).
98. Zhang, N., Duncan, F. E., Que, E. L., O'Halloran, T. V. & Woodruff, T. K. The fertilization-induced zinc spark is a novel biomarker of mouse embryo quality and early development. *Sci. Rep.* **6,** 22772 (2016).
99. Liu, D. Y. & Baker, H. W. G. Defective sperm–zona pellucida interaction: a major cause of failure of fertilization in clinical in-vitro fertilization. *Hum. Reprod.* **15,** 702–708 (2000).
100. Margalit, M. *et al.* Genetic and physiological study of morphologically abnormal human zona pellucida. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **165,** 70–76 (2012).
101. Ferré, M. *et al.* Are zona pellucida genes involved in recurrent oocyte lysis observed during in vitro fertilization? *J. Assist. Reprod. Genet.* **31,** 221–227 (2014).
102. Männikkö, M. *et al.* Association between sequence variations in genes encoding human zona pellucida glycoproteins and fertilization failure in IVF. *Hum. Reprod.* **20,** 1578–1585 (2005).
103. Yang, P. *et al.* Novel zona pellucida gene variants identified in patients with oocyte anomalies. *Fertil. Steril.* **107,** 1364–1369 (2017).
104. Huang, H.-L. *et al.* Mutant ZP1 in familial infertility. *N. Engl. J. Med.* **370,** 1220–1226 (2014).
105. Barbaux, S., El Khattabi, L. & Ziyyat, A. ZP2 heterozygous mutation in an infertile woman. *Hum. Genet.* **136,** 1489–1491 (2017).
106. Chen, T. *et al.* A Recurrent Missense Mutation in ZP3 Causes Empty Follicle Syndrome and Female Infertility. *Am. J. Hum. Genet.* **101,** 459–465 (2017).
107. Dai, C. *et al.* ZP2 pathogenic variants cause in vitro fertilization failure and female infertility. *Genet. Med.* (2018). doi:10.1038/s41436-018-0064-y
108. Chandra, A., Copen, C. E. & Mosher, W. D. Sexual Behavior, Sexual Attraction, and Sexual Identity in the United States: Data from the 2006–2010 National Survey of Family Growth. in *International Handbook on the Demography of Sexuality* (ed. Baumle, A. K.) 45–66 (Springer Netherlands, 2013).
109. Organization, W. H. & Others. Improving access to quality care in family planning: medical eligibility criteria for contraceptive use. (2000).
110. Li, J. *et al.* Vectored antibody gene delivery mediates long-term contraception. *Curr. Biol.* **25,** R820–2 (2015).
111. Shivers, C. A., Dudkiewicz, A. B., Franklin, L. E. & Fussell, E. N. Inhibition of sperm-egg interaction by specific antibody. *Science* **178,** 1211–1213 (1972).
112. Jílek, F. & Pavlok, A. Antibodies against mouse ovaries and their effect on fertilization in vitro and in vivo in the mouse. *J. Reprod. Fertil.* **42,** 377–380 (1975).
113. Oikawa, T. & Yanagimachi, R. Block of hamster fertilization by anti-ovary antibody. *J. Reprod. Fertil.* **45,** 487–494 (1975).
114. Sacco, A. G. Antigenic cross-reactivity between human and pig zona pellucida. *Biol. Reprod.* **16,** 164–173 (1977).
115. Wood, D. M., Liu, C. & Dunbar, B. S. Effect of alloimmunization and heteroimmunization with zonae pellucidae on fertility in rabbits. *Biol. Reprod.* **25,** 439–450 (1981).
116. Kirkpatrick, J. F., Liu, I. M., Turner, J. W., Jr, Naugle, R. & Keiper, R. Long-term effects of porcine zonae pellucidae immunocontraception on ovarian function in feral horses (Equus caballus). *J. Reprod. Fertil.* **94,** 437–444 (1992).
117. Mahi-Brown, C. A. & Huang, T. T. F., Jr. Infertility in bitches induced by active immunization with porcine zonae pellucidae. *Journal of* (1982).
118. Sacco, A. G., Yurewicz, E. C., Subramanian, M. G., Lian, Y. & Dukelow, W. R. Immunological response and ovarian histology of squirrel monkeys (Saimiri sciureus) immunized with porcine zona pellucida ZP3 (Mr = 55,000) macromolecules. *Am. J. Primatol.* **24,** 15–28 (1991).
119. Kirkpatrick, J. F., Turner, J. W., Jr, Liu, I. K. & Fayrer-Hosken, R. Applications of pig zona pellucida immunocontraception to wildlife fertility control. *J. Reprod. Fertil. Suppl.* **50,** 183–189 (1996).
120. Fayrer-Hosken, R. A., Grobler, D., Van Altena, J. J., Bertschinger, H. J. & Kirkpatrick, J. F. Immunocontraception of African elephants. *Nature* **407,** 149 (2000).
121. Hardy, C. M., Beaton, S. & Hinds, L. A. Immunocontraception in mice using repeated, multi-antigen peptides: immunization with purified recombinant antigens. *Mol. Reprod. Dev.* **75,** 126–135 (2008).
122. Martinez, M. L. & Harris, J. D. Effectiveness of zona pellucida protein ZPB as an immunocontraceptive antigen. *J. Reprod. Fertil.* **120,** 19–32 (2000).
123. Dean, J. & East, I. J. Effects of Anti-Zona Pellucida Monoclonal Antibodies on Fertilization and Early Development. in *The Molecular and Cellular Biology of Fertilization* (ed. Hedrick, J. L.) 37–53 (Springer US, 1986).
124. Millar, S. E. *et al.* Vaccination with a synthetic zona pellucida peptide produces long-term contraception in

female mice. *Science* **246,** 935–938 (1989).
125. Rhim, S. H. *et al.* Autoimmune disease of the ovary induced by a ZP3 peptide from the mouse zona pellucida. *J. Clin. Invest.* **89,** 28–35 (1992).
126. Joonè, C. J., Schulman, M. L. & Bertschinger, H. J. Ovarian dysfunction associated with zona pellucida-based immunocontraceptive vaccines. *Theriogenology* **89,** 329–337 (2017).
127. Thaler, C. D. & Cardullo, R. A. The initial molecular interaction between mouse sperm and the zona pellucida is a complex binding event. *J. Biol. Chem.* **271,** 23289–23297 (1996).
128. Wright, G. J. & Bianchi, E. The challenges involved in elucidating the molecular basis of sperm–egg recognition in mammals and approaches to overcome them. *Cell Tissue Res.* **363,** 227–235 (2016).
129. Nooren, I. M. A. & Thornton, J. M. Structural Characterisation and Functional Significance of Transient Protein–Protein Interactions. *J. Mol. Biol.* **325,** 991–1018 (2003).
130. Malashkevich, V. N., Kammerer, R. A., Efimov, V. P., Schulthess, T. & Engel, J. The crystal structure of a five-stranded coiled coil in COMP: a prototype ion channel? *Science* **274,** 761–765 (1996).
131. Voulgaraki, D. *et al.* Multivalent recombinant proteins for probing functions of leucocyte surface proteins such as the CD200 receptor. *Immunology* **115,** 337–346 (2005).
132. van der Merwe, P. A., Brown, M. H., Davis, S. J. & Barclay, A. N. Affinity and kinetic analysis of the interaction of the cell adhesion molecules rat CD2 and CD48. *EMBO J.* **12,** 4945–4954 (1993).
133. Corzo, J. Time, the forgotten dimension of ligand binding teaching. *Biochem. Mol. Biol. Educ.* **34,** 413–416 (2006).
134. Wang, H.-W. & Wang, J.-W. How cryo-electron microscopy and X-ray crystallography complement each other. *Protein Sci.* **26,** 32–39 (2017).
135. Vergara, G. J., Irwin, M. H., Moffatt, R. J. & Pinkert, C. A. In vitro fertilization in mice: Strain differences in response to superovulation protocols and effect of cumulus cell removal. *Theriogenology* **47,** 1245–1252 (1997).
136. Luo, C. *et al.* Superovulation strategies for 6 commonly used mouse strains. *J. Am. Assoc. Lab. Anim. Sci.* **50,** 471–478 (2011).
137. Litscher, E. S. & Wassarman, P. M. Recombinant hamster sperm receptors that exhibit species-specific binding to sperm. *Zygote* **4,** 229–236 (1996).
138. Muro, Y., Buffone, M. G., Okabe, M. & Gerton, G. L. Function of the acrosomal matrix: zona pellucida 3 receptor (ZP3R/sp56) is not essential for mouse fertilization. *Biol. Reprod.* **86,** 1–6 (2012).
139. Strub, M.-P. *et al.* Selenomethionine and Selenocysteine Double Labeling Strategy for Crystallographic Phasing. *Structure* **11,** 1359–1367 (2003).
140. Barton, W. A., Tzvetkova-Robev, D., Erdjument-Bromage, H., Tempst, P. & Nikolov, D. B. Highly efficient selenomethionine labeling of recombinant proteins produced in mammalian cells. *Protein Sci.* **15,** 2008–2013 (2006).
141. Silkowski, H. *et al.* Characterisation of the low affinity interaction between rat cell adhesion molecules CD2 and CD48 by analytical ultracentrifugation. *Eur. Biophys. J.* **25,** 455–462 (1997).
142. Isotani, A. *et al.* A delayed sperm penetration of cumulus layers by disruption of acrosin gene in rats. *Biol. Reprod.* **97,** 61–68 (2017).
143. Mortillo, S., Stewart, C. L. & Wassarman, P. M. Embryonal carcinoma cells transfected with ZP3 genes differentially glycosylate similar polypeptides and secrete active mouse sperm receptor. *J. Cell Biol.* (1991).
144. Algarra, B. *et al.* Deletion of the C-terminal region of OVGP1 affects porcine IVF. *RA* (2014). doi:10.1530/repabs.1.P331

145. link: http://www.dnastar.com/protean3d_help/#!Documents/listofproteases.htm